

Bügelmayer, Elisabeth; Schnitzlein, Daniel D.

Working Paper

Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health

SOEPpapers on Multidisciplinary Panel Data Research, No. 716

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Bügelmayer, Elisabeth; Schnitzlein, Daniel D. (2014) : Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health, SOEPpapers on Multidisciplinary Panel Data Research, No. 716, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/106180>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SOEPpapers

on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel Study at DIW Berlin

716-2014

Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health

Elisabeth Bügelmayer and Daniel D. Schnitzlein

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
<http://www.diw.de/soeppapers>

Editors:

Jürgen **Schupp** (Sociology)

Gert G. **Wagner** (Social Sciences, Vice Dean DIW Graduate Center)

Conchita **D'Ambrosio** (Public Economics)

Denis **Gerstorff** (Psychology, DIW Research Director)

Elke **Holst** (Gender Studies, DIW Research Director)

Frauke **Kreuter** (Survey Methodology, DIW Research Professor)

Martin **Kroh** (Political Science and Survey Methodology)

Frieder R. **Lang** (Psychology, DIW Research Professor)

Henning **Lohmann** (Sociology, DIW Research Professor)

Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Professor)

Thomas **Siedler** (Empirical Economics)

C. Katharina **Spieß** (Empirical Economics and Educational Science)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

Is it the family or the neighborhood?

Evidence from sibling and neighbor correlations in youth education and health

Elisabeth Bügelmayer
DIW Berlin

Daniel D. Schnitzlein
Leibniz University Hannover, DIW Berlin

This version: December 2014

Abstract

In this paper we present sibling and neighbor correlations in school grades and cognitive skills as well as indicators of physical and mental health for a sample of German adolescents. In a first step, we estimate sibling correlations and find substantial influence of shared family and community background on all outcomes. To further disentangle the influence of family background and neighborhood, we estimate neighbor correlations. Our results show that for all outcomes, estimated neighbor correlations are clearly lower than estimated sibling correlations. However, especially for cognitive skills and mental health, neighbor correlations are still substantial in relation to sibling correlations. Thus, compared to existing results from other countries, the influence of the neighborhood is not negligible in Germany for these outcomes.

JEL codes: J62

Keywords: Sibling correlations, intergenerational mobility, neighbor correlations.

Correspondence to:

Elisabeth Bügelmayer, DIW Berlin, SOEP Department, Mohrenstr. 58, 10117 Berlin, Germany, Tel.: +49 30 89789-344, Fax: +49 30 89789-200, E-mail: ebuegelmayer@diw.de

Acknowledgments:

We thank Karin Hederos Eriksson and participants of ESPE 2014, EALE 2014, and workshop participants at RWI Essen for helpful comments and suggestions. Elisabeth Bügelmayer gratefully acknowledges support from the Leibniz Association, Bonn, under the grant “Neighborhood effects: Analysis of individual rational behaviour in a social context.”

We must not tolerate a situation where children cannot develop their talents because there is no equality of opportunity. We must not tolerate a situation where people have the impression that there is no longer any point in putting in any effort because they won't progress even if they work hard.
(Gauck, 2012)

1 Introduction

In his 2012 inauguration speech, German President Joachim Gauck formulated equality of opportunity as a normative policy goal to be reached. This normative goal, which is the foundation of most Western societies, not just Germany, implies that an individual's success is independent of factors beyond his or her control (Roemer, 1998).

The family and the neighborhood in which a child grows up are probably the two most important examples of such factors. Families influence their children in numerous ways, including the provision of resources, transmission of characteristics and skills, as well as investments in skill formation.¹ The neighborhood a child grows up in might influence child outcomes through various channels, such as social contagion and networks, environmental factors, as well as access to public services and infrastructure.²

In the economic literature the traditional approach to analyze equality of opportunity in a society is to estimate measures of intergenerational mobility – for example intergenerational correlations or intergenerational elasticities. This line of research has been extended to the analysis of sibling correlations (see for example Solon et al., 1991, Solon, 1999, Björklund and Jäntti, 2012). In contrast to intergenerational correlations – which only cover bivariate relationships – sibling correlations are a much broader measure of the influence of family and community background. In particular, sibling correlations measure the importance of all factors that siblings of one family share for their later outcomes. As this includes both, family background characteristics and neighborhood factors, sibling

¹ See for example Cunha and Heckman (2007, 2008), Currie (2009), Anger and Heineck (2010), Andrabi et al. (2012), Carneiro et al. (2013).

² See Galster (2012) and Durlauf (2004) for an overview of the literature on neighborhood effects.

correlations can be seen as a measure of equality of opportunity as defined above.³ The higher the similarity between siblings, the more important are family and neighborhood characteristics and the lower is the level of equality of opportunity.

Authors estimate sibling correlations, among others, in monetary outcomes (e.g., Solon et al., 1991, Björklund et al., 2002, Mazumder, 2008, Schnitzlein, 2014), educational outcomes (Björklund and Salvanes, 2011, Nicoletti and Rabe, 2013), and health outcomes (Mazumder, 2011) for the US, UK, Germany and Scandinavian countries. The results show substantial influence of shared family and neighborhood background over all outcomes, but also significant cross-country differences in the importance of the combined effect.

The important question in interpreting these results is which factor is the main determinant? Or in other words, *is it the family or the neighborhood?* To address this question, analogous to sibling correlations, neighbor correlations are estimated. They provide a descriptive summary measure of how much of the outcome of neighbors can be attributed to the shared neighborhood. Solon et al. (2000), Page and Solon (2003), Raaum et al. (2006), Lindahl (2011) and Nicoletti and Rabe (2013) provide evidence for the US, Norway, Sweden and the UK, finding only weak neighbor correlations compared to sibling correlations for monetary and education outcomes.

In this paper, we focus on two crucial components of human capital formation, namely education and health. More specifically, we analyze school grades, cognitive skills, BMI, mental health, physical health, and height of a sample of German adolescents.

Regarding education, there are several studies investigating family and neighbor correlations. Solon et al. (2000) apply the method of sibling correlations to neighboring children. Their sample comprises 687 individuals from 379 families from the PSID. Sibling correlations in years of education are estimated to be 0.5. The correlation in educational

³ Note that sibling correlations measure only factors that are shared by siblings. Therefore, a sibling correlation is still a lower bound measure of the true influence of the family (Björklund and Jäntti, 2012). For more details, see section 3.

attainment of neighboring children is 0.10 when controlling for basic family characteristics. Therefore, Solon et al. (2000) conclude that inequalities in educational attainment are mainly attributed to family background rather than growing up in the same neighborhood. Raaum et al. (2006) use census data from Norway to analyze the role of family and neighborhood on adult education and earnings. They report neighbor correlations in years of schooling of around 0.10. To adjust for the fact that similarities in adult outcomes might only be due to the fact that children growing up in the same neighborhood share similar family characteristics, again family background is controlled for. However, when adjusting for family background (parental education, family structure and parental income), neighbor correlations fall to around 0.04. Thus, childhood neighborhoods are substantially less important for adult outcomes than families. Lindahl (2011) assesses the importance of family versus neighborhood factors. She estimates correlations in income and education among siblings and children from the same neighborhood for about 13,000 individuals in Sweden. Applying a two-level model she finds that about 40% of the variation between siblings in years of education is due to shared background factors. Correlations among neighboring youth make up 15-20% of the sibling correlation. When family background is controlled for, neighborhood correlations drop to less than 3%. Lindahl (2011) concludes that family background is substantially more important than the neighborhood. This result is confirmed by Nicoletti and Rabe (2013) who also apply multilevel models to estimate the influence of family and neighborhood factors on pupil's test scores in the UK. At age 16, they report magnitudes of sibling and neighbor correlations of 0.61 and 0.14, respectively. These figures are comparable to the ones estimated by Solon et al. (2000) and by Raaum et al. (2006).

With respect to health, there is little evidence on sibling correlations. One exception is Mazumder (2011). He estimates sibling correlations in health using the PSID. Results show that correlations in siblings' outcomes are already high at birth and do not decline

significantly over the life span. Even less research exists on neighbor correlations in health, although neighborhoods possibly affect health through physical neighborhood conditions such as pollution, poor quality housing, and stress of living in a dangerous neighborhood (see Robert, 1998). Further, the provision of medical infrastructure differs across neighborhoods.⁴

The empirical evidence for both education and health indicators in Germany is scarce. While Sieben et al. (2001) and Schnitzlein (2014) present sibling correlations in educational outcomes, neighbor correlations in education outcomes as well as sibling and neighbor correlations in health outcomes have not been analyzed for Germany so far. We focus on a sample of adolescents as they are often perceived as being most susceptible to disadvantageous neighborhood impacts (Kling et al., 2007). Therefore it is important to analyze the level of equality of opportunity this group faces. This, in turn, allows for a comparison of the extent of inequality in Germany versus other countries.

The contributions of our paper to the literature are threefold. First, we estimate sibling correlations in youth health and education and thus provide novel evidence on the joint importance of family background and neighborhood influence for youth outcomes in Germany. Second, we contrast the sibling correlations with estimated neighbor correlations in these outcomes. This enables us to give a first answer to the question of whether it is the family or the neighborhood that matters most in Germany. Third, we try to unveil the reasons and channels at work behind the neighborhood influence in Germany we find. To this end, we combine our survey data on adolescents with small scale neighborhood indicators from administrative register data. This data provides us with information about the average

⁴ Some evidence on the neighborhood impact on mental and physical health is provided by the Moving to Opportunity (MTO) study, a large randomized housing mobility experiment. Leventhal and Brooks-Gunn (2003) show that children benefitted most in terms of mental health when moving to a better neighborhood. Kling et al. (2007) find that teenage girls largely benefit in terms of mental health while teenage boys experience a decrease in these measures after moving to a lower poverty neighborhood. In the domain of physical health, they find no effect on adults (except obesity) and small effects on adolescents' physical health. Also using MTO data, Fortson and Sanbonmatsu (2010) focus on the influence of neighborhood quality on the physical health of children between 6 and 20 years of age. However, they do not find any positive impact on self-rated health, BMI, asthma or injuries in the medium-term.

economic situation and the average level of education within a German postal code area and has rarely been employed so far.⁵ We use this information to decompose the neighbor correlations with respect to neighborhood characteristics in order to investigate which pathways explain the neighborhood influence on our health and education outcomes.

We find substantial sibling correlations in youth education and health, albeit some of our estimated sibling correlations are lower than estimates for other countries. In line with the previous literature, our neighbor correlations are clearly lower than the estimated sibling correlations. However, our estimates show that for some youth outcomes, such as cognitive skills and mental health, the neighbor correlations still are substantial in relation to the sibling correlations. After controlling for parental characteristics, neither economic indicators nor the level of education in the neighborhood seem to be important determinants of these correlations.

The reminder of the paper is organized as follows. Section 2 discusses the empirical strategy. Section 3 discusses our data and our outcome variables. Section 4 contains our results and a discussion and section 5 concludes.

2 Econometric model and empirical strategy

Studies analyzing the neighborhood impact on children's outcomes face several econometric challenges due to non-random sorting of families into neighborhoods. Neighborhood characteristics and family characteristics are highly correlated as families self-select into neighborhoods. Due to this self-selection, simply regressing child outcomes on a set of family characteristics and neighborhood indicators makes it very difficult to disentangle the family effect from the neighborhood effect.⁶ Solon et al. (2000) were the first to apply the method of

⁵Exceptions are Bauer et al. (2012), Bauer et al. (2013), and Hawranek and Schanne (2014).

⁶In an overview of numerous studies applying this approach, Ginther et al. (2000) find a very large variation in results depending on the choice of the control variables and the selection of the neighborhood characteristics.

sibling correlations to neighboring children as a solution to this problem: in an ideal world we would like to estimate the following model for the relationship between a child's (index i) outcome of interest y_{nfi} and his family and neighborhood characteristics as

$$y_{nfi} = \alpha' X_{nf} + \beta' Z_n + \varepsilon_{nfi} \quad (1)$$

with X_{nf} being a matrix of family (index f) variables, Z_n being a matrix of neighborhood variables (index n), α' and β' are the parameters to be estimated, with an error term, ε_{nfi} , that is uncorrelated with both family and neighborhood characteristics.

Based on the assumption that similar families tend to sort into similar neighborhoods (Tiebout, 1956), we expect X_{nf} and Z_n to be positively correlated. To obtain unbiased estimates of α and β from equation (1), ε_{nfi} has to be uncorrelated with both family and neighborhood factors. This only holds if we assume that X_{nf} and Z_n fully cover all relevant family and neighborhood information. As this is most certainly not the case, a simple estimation of equation (1) cannot answer the question raised in the introduction if family background or neighborhood factors are the most important determinant of a child's outcome.

We follow Solon et al. (2000) and adopt the idea raised in the literature on sibling correlations to estimate the importance of the neighborhood via the similarity of two individuals with identical community background. In particular, we estimate sibling and neighbor correlations in y_{nfi} . To compute these correlations, we first need an expression for the population variance of y_{nfi} . If we assume the model in equation (1), the population variance of y_{nfi} is the following

$$Var(y_{nfi}) = Var(\alpha' X_{nf}) + Var(\beta' Z_n) + 2Cov(\alpha' X_{nf}, \beta' Z_n) + Var(\varepsilon_{nfi}) \quad (2)$$

where $Var(\alpha'X_{nf})$ is the variance in family characteristics and $Var(\beta'Z_n)$ is the variance in neighborhood characteristics. These can be seen as the genuine family and neighborhood effects (Lindahl, 2011). $Cov(\alpha'X_{nf}, \beta'Z_n)$ instead measures the covariance between family and neighborhood characteristics. Due to the nonrandom sorting, this covariance is assumed to be positive.

The covariance in y_{nfi} of two siblings, i and i' , from the same family equals

$$Cov(y_{nfi}, y_{nfi'}) = Var(\alpha'X_{nf}) + Var(\beta'Z_n) + 2Cov(\alpha'X_{nf}, \beta'Z_n) \quad (3)$$

and the covariance in y_{nfi} of two neighboring children, i and i' , from different families f and f' in the same neighborhood equals

$$Cov(y_{nfi}, y_{nf'i'}) = Cov(\alpha'X_{nf}, \alpha'X_{nf'}) + Var(\beta'Z_n) + 2Cov(\alpha'X_{nf}, \beta'Z_n). \quad (4)$$

Following the argumentation in Solon et al. (2000), we expect $Cov(\alpha'X_{nf}, \alpha'X_{nf'})$ – the covariance between family characteristics in the same neighborhood – to be positive because of self-selection of families into neighborhoods. At the same time we expect $Var(\alpha'X_{nf})$ to be larger than $Cov(\alpha'X_{nf}, \alpha'X_{nf'})$ because children from the same family are more similar than children from different families. Therefore, we expect the neighbor correlation to be smaller than the sibling correlation:

$$\rho_{neighbor} = \frac{Cov(y_{nfi}, y_{nf'i'})}{Var(y_{nfi})} < \rho_{sibling} = \frac{Cov(y_{nfi}, y_{nfi'})}{Var(y_{nfi})} \quad (5)$$

Following Solon et al. (1991), we estimate the sibling and the neighbor correlation as intra-class correlations in a linear multilevel model. To illustrate this, suppose we observe an outcome y_{ig} for an individual i from group g . In this formulation, group can either be the family or the neighborhood. We now assume that y_{ig} can be modeled as

$$y_{ig} = \alpha_g + \mu_{ig} \quad (6)$$

where α_g is a group (family/neighborhood) specific component and μ_{ig} is an individual specific component. The group specific component covers all factors that are common to the individuals from the same group. In the case of siblings, these include not just family background factors, but also neighborhood background factors, and, in the case of neighbors, these include only neighborhood characteristics. With one individual being observed only in one group, these two components are orthogonal to each other. We therefore can write the variance of the outcome as sum of the variances of the two components:

$$Var(y) = Var(\alpha) + Var(\mu) \quad . \quad (7)$$

The correlation ρ between the outcome of two individuals, i and i' , of one group then can be calculated as the ratio of the variance of the group specific component and the overall variance of the outcome:

$$\rho = \frac{Var(\alpha)}{Var(\alpha) + Var(\mu)} \quad (8)$$

Following Mazumder (2008), these variance components can be estimated via restricted maximum likelihood in the following model:

$$y_{ig} = \beta X_{ig} + \alpha_g + \mu_{ig} \quad (9)$$

where y_{ig} is the observed outcome of individual i in family/neighborhood g . X is a matrix of control variables containing year dummies, α_g is the family/neighborhood component, and μ_{ig} is the individual component. The standard errors of the sibling and neighborhood correlations are calculated via the bivariate delta method.

As can be seen from equation (4), the raw neighbor correlation covers the neighborhood effect as well as the effect of positive sorting of families into neighborhoods. Therefore a raw estimate of a neighborhood correlation would overestimate the importance of the neighborhood. To – at least partly – solve this problem, we present neighbor correlations adjusted for family background influences to account for the fact that similarities in neighboring adolescent's outcomes might only be due to similar family characteristics. We include family background variables in X_{ig} in equation (9) (Solon et al., 2000, Raaum et al., 2006, Lindahl, 2011). This – adjusted – neighbor correlation, which is still an upper bound estimate since we do not observe all relevant family characteristics, can now be compared to the sibling correlation to answer the question raised in the introduction.

In a next step, we add neighborhood characteristics to X_{ig} . This approach is suggested by Mazumder (2008) in order to uncover channels through which the influence of family background works. We adapt this approach for neighborhoods. If the added characteristics have an effect in determining the outcome y_{ig} , this will reduce the variance of the neighborhood component in equation (8), and thus reduce the neighbor correlation. The size of this reduction can be interpreted as a measure of the importance of this specific influence factor.

3 Data

Our study is based on German survey and administrative data. Our main data source is the German Socio-Economic Panel Study (SOEP). The SOEP provides detailed information on adolescents' health and educational attainment. It is an annual representative, longitudinal household panel with about 11,000 participating households and more than 20,000 individuals in the recent wave (Wagner et al., 2007). Since 2000, adolescents at age 17 fill in a youth specific questionnaire. Among other information, this questionnaire obtains detailed information about the youth's health and school performance.

3.1 Youth education and health outcomes

The SOEP data contains rich information on adolescents' educational performance. Specifically, as indicators of school performance, grades for both German and math of the last school report card are available. In Germany, grades range between 1 and 6, with 6 being the worst grade. In our estimations, we also include the average of these two grades as a combined measure of school performance.

Since 2006, the SOEP youth questionnaire includes a measure of cognitive ability. It consists of three parts, which were taken from the I-S-T 2000 (Solga et al., 2005), namely verbal skills, numerical skills, and abstract reasoning. The first two are measures of crystalline intelligence and the last measures fluid intelligence, i.e. innate ability (Anger, 2012, Richter et al., 2013). We use the sum of correct answers in all three tests as our measure of cognitive skills.

Self-reported height and weight, which we use to calculate the body mass index (BMI) for each individual, is also included in the youth questionnaire. As measures for physical and mental health we use the SF12v2 inventory, which, since 2002, is asked every second year in the SOEP. The SF12v2 contains twelve health questions, which are converted into continuous

subscales of mental and physical health by a special algorithm (Andersen et al., 2007). As the SF12v2 questions are not available in the youth questionnaire, we use averages of the subsequent years in mental health status and physical health status in which the participants answer the regular SOEP questionnaire.

Furthermore, we include height as sole outcome variable in our analysis. Height should be mainly affected by genetic factors and therefore family background, but should be largely independent of the neighborhood (Duncan et al., 2005). It also serves as an indicator of comparableness of our results with existing studies. Table 1 shows descriptive statistics of our outcome variables.

3.2 Family and neighborhood characteristics

As discussed in the last section, raw neighborhood correlations would overestimate the influence of the neighborhood and, therefore, have to be adjusted by including family characteristics. Thus, we include the parental household income and the highest education achieved by either parent when the adolescents are 17 years old. Both indicators are available in the SOEP data. Table 2 shows descriptive statistics of these parental characteristics.

For a subset of our observations, we have detailed neighborhood information available. This information on the neighborhoods is mainly based on administrative data collected by the Institute for Employment Research (IAB) in Nuremberg. Specifically, our indicators, which are drawn from the official employment and unemployment registers, are available for the 2004 to 2010 period at the postal code level.⁷ We merge our SOEP data with these administrative data on neighborhood characteristics. Available characteristics are the

⁷ The research data center of the Federal Employment Agency at the Institute for Employment Research provides the SIAB (vom Berge et al., 2013), which is a sample of these data and can be used either at the RDC or can also be obtained as a scientific use file. However, our neighborhood indicators are calculated from the full data, the Integrated Employment Biographies (IEB), not the sample.

share of workers with a university degree in the workforce⁸ and the share of workers with a high school diploma in the workforce, which we use as indicators for the education level in the neighborhood. As indicators for the economic situation of the neighborhood, we include the share of unemployed in the workforce. In addition, we add the share of indebted people in the population over 18, which is provided by one of the largest private debt collection enterprises in Germany and is available for use with the SOEP data at DIW Berlin. Table 3 shows descriptive statistics of the neighborhood variables. Overall, they display substantial variation across postal code areas.⁹

4 Results

4.1 Sibling correlations in youth education and health

We begin the discussion of the results with the estimated sibling correlations. In Table 4, the estimated sibling correlations are shown alongside the estimated variance components as well as the number of families and observations. The sibling correlations in school grades in Germany are lower than expected, based on results from other countries. Our estimates of sibling correlations range from 0.17-0.23 for the German grade, the math grade, and the average of both. That means family and community background together explain between 17 and 23 percent of the variation in grades. Nicoletti and Rabe (2013) find sibling correlations in test scores of around 0.6 for the UK, while Mazumder (2011) estimates sibling correlations

⁸ The workforce was computed as the sum of employed persons who are subject to social insurance contributions and registered unemployed persons (Bauer et al., 2011). Registered unemployed persons also include persons looking for work and persons who are liable to social security assistance if a family member is unemployed. People who are looking for work and are not registered as unemployed, self-employed and civil servants are not captured.

⁹ Missing observations originate from data protection, which forbids reporting indicators for a postal code areas with fewer than 20 registered persons in the official administrative statistics.

in high school GPAs of about 0.3 for the US. So, for our sample, the influence of family and community background on school performance is less pronounced than in the UK or the US.¹⁰

Further, we estimate a sibling correlation in cognitive ability of 0.46. So about half of the variation in cognitive ability can be attributed to factors shared by siblings. This estimate is comparable in size to studies using Swedish administrative data (Björklund et al., 2010, Björklund and Jäntti, 2012). Mazumder (2008) reports – based on data from the National Longitudinal Survey of Youth (NLSY) – only slightly higher correlations for measures of cognitive skills in the US.

The sibling correlations in our mental health and physical health outcomes range between 0.31 and 0.21, respectively. These results are lower than the 0.43 (mental health) and 0.37 (general health status) that Mazumder (2011) reports. However, health measures and age ranges are not directly comparable with Mazumder (2011). For BMI and height we estimate sibling correlations of 0.33 and 0.42. Concerning BMI, our estimates are comparable with findings from Mazumder (2008), who reports sibling correlations around 0.30. Regarding height, this is in line with the results Schnitzlein (2014) finds in an adult German sample and Duncan et al. (2001) find for adolescents in the US.

4.2 Neighbor correlations in youth education and health

Table 5 contains the corresponding neighbor correlations. The upper panel (Panel A) shows the raw *unadjusted* neighbor correlations and the lower panel (Panel B) shows neighbor correlations *adjusted* for family characteristics (parental education and household income) as discussed in section 2 to account for the possibility that outcomes of neighboring youth resemble each other not because of growing up in the same neighborhood but because they share the same family characteristics (see e.g., Raaum et al., 2006).

¹⁰ Note that grades are not directly comparable to test scores and might be a noisier measure than the test scores used in the previous literature.

We find that the shared neighborhood accounts for around 8 percent of the variance in grades. Adjusting for parental background reduces the neighbor correlations to 0.06 - 0.08. So only between 6 and 8 percent of the variation in school grades can be attributed to the neighborhood.¹¹ In total numbers, this is even lower than the 0.14 Nicoletti and Rabe (2013) report for the test scores of 16 year olds in the UK. If we compare the sibling correlation to the neighbor correlation, the relation of both is roughly in the same range in the UK and in Germany.

Notably, neighbor correlations in cognitive ability are considerably higher and amount to 0.27. When family characteristics are controlled for, this correlation declines to 0.22. Still, about 22 percent of the variation in cognitive ability test scores can be attributed to the neighborhood.

Concerning health, the highest neighbor correlation we estimate is in mental health. Shared neighborhood factors at age 17 account for 14 percent of the variance in mental health. In contrast, neighbor correlations in physical health are around 0.09; adjusting for parental background leaves neighbor correlations in mental health and physical health largely unaffected. Adjusted neighbor correlations in BMI amount to 0.10. Concerning height, we estimate an adjusted neighbor correlation of 0.15, which is considerably higher than the 1% Duncan et al. (2001) find.

This relatively high neighbor correlation in height seems suspect, unless one assumes sorting into neighborhoods based on height or neighborhood conditions that strongly affect inhabitants' height. We test the robustness of our results to rule out the possibility that our neighbor correlations are simply artifacts of our sibling correlations in neighborhoods with a small number of (large) families. To this end, we restrict our adjusted neighbor correlations to youth who live in neighborhoods in which we observe more than two families in the data. The

¹¹ Note that – as argued in section 3 – this is still an upper bound estimate of the influence of the neighborhood.

results are presented in Table 6. Comparing these results to Panel B in Table 5 shows that neighbor correlations in height (and BMI) drop substantially, becoming insignificant. The other neighbor correlations slightly decline, except for cognitive ability. Thus, the substantial neighborhood influence on cognitive ability and mental health proves to be a robust result.

4.3 Comparison between sibling and neighbor correlations

Finally, we combine these results in Figure 1 to shed light on our question raised in the introduction. *Is it the family or the neighborhood?* The estimated sibling correlations, along with the adjusted neighbor correlations and the neighbor correlations restricted to youth living in neighborhoods with more than two families, are presented.

The first result is, as implied in equation (5), that the estimated neighbor correlations are all significant, but smaller in size compared to the estimated sibling correlations.¹² This result holds for all outcomes. If we compare the sibling correlations and the adjusted neighbor correlations, for all outcomes except cognitive ability and mental health, neighbor correlations are about one-third of the estimated sibling correlation. Adding the results from our robustness test shows a slightly increase in the difference between sibling and neighbor correlations. So the answer to the question is that it is *mostly* the family that influences the adolescent outcomes, with a *minor* part attributable to the neighborhood. Notably, for cognitive ability and mental health, the influence of the neighborhood is higher.

If we compare our results with existing results from other countries (Solon et al., 2000, Raaum et al., 2006, Lindahl, 2011, Nicoletti and Rabe, 2013) – although family influence is most important – neighborhood influence is more important in relation to family influence in Germany. Thus, the results from the existing literature do not fully carry over to the German context.

¹² Technically, equation (5) refers to the unadjusted neighbor correlation, but the result also holds for the adjusted neighbor correlation.

In a last step, we aim at explaining channels through which the neighbor correlations we find might work. In a subsample of our youth data, we have detailed neighborhood characteristics available, including the level of education and the economic situation in the neighborhood. Following the decomposition approach by Mazumder (2008), we add these indicators to the estimation of the neighbor correlations. Further, we include community size as Page and Solon (2003) and Nicoletti and Rabe (2013) find significant effects of urbanicity on wages and test scores. This approach aims at disentangling the impact of shared family factors from shared community factors and is an extension of the previous literature, which often lacks joint information on family and neighborhood background.

The results are presented in Table 7 and show that adding these neighborhood characteristics does not decrease neighbor correlations after controlling for parental characteristics.¹³ Thus, the educational structure and the economic situation in the neighborhood do not additionally contribute to the explanation of the influence of the neighborhood.

5 Conclusion

In this paper we present sibling and neighbor correlations in education and health for a sample of German 17 year olds. Moreover, we add to the literature by providing evidence on neighbor correlations in health outcomes, which, so far, have been scarcely investigated. Our analysis enables a cross-country comparison of the extent of inequality, which can be ascribed to growing up in the same family or in the same neighborhood in Germany.

Overall, we find evidence of substantial joint influence of family and community background on youth school grades, cognitive skills, mental health, physical health, BMI, and height. The estimated sibling correlations are partly lower than comparable estimates for other

¹³ Some of the neighbor correlations adjusted to parental characteristics are higher in the subsample than in our full sample. This is most likely due to the reduced sample as neighborhood characteristics are only available for the years 2004-2010. We estimated various specifications, including neighborhood variables sequentially or simultaneously but results remained nearly unchanged. Estimations are available upon request.

countries, but are, consistent with the literature, always larger than the neighbor correlations. However, compared to existing results from other countries, the influence of the neighborhood is not negligible in Germany for some outcomes. This applies especially to the domains of cognitive ability and mental health, suggesting that, next to indicators of economic self-sufficiency, further research on the neighborhood impact to individual health is needed.

Our results have important implications for policy makers. While the existing literature emphasizes the importance of support for children with adverse family backgrounds, our findings suggest that neighborhood quality should also be a target of equality enhancing policies.

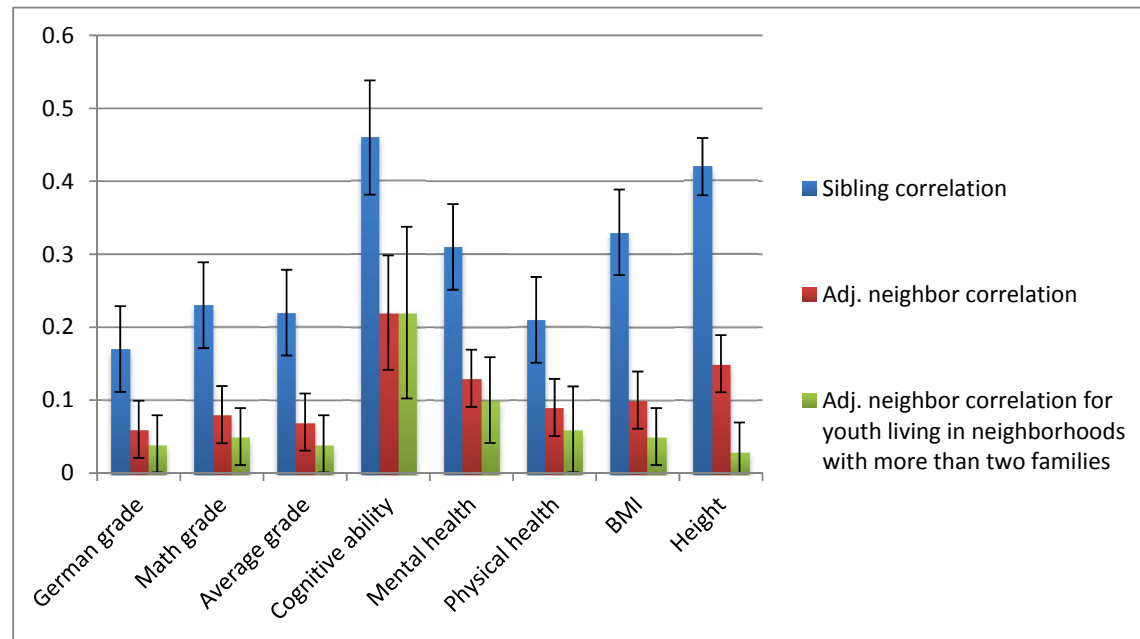
References

- Andersen, H.H., A. Mühlbacher, M. Nübling, J. Schupp and G.G. Wagner, "Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of Sf-12v2," *Schmollers Jahrbuch*, 127(1), 171-182, 2007.
- Andrabi, T., J. Das and A.I. Khwaja, "What Did You Do All Day? Maternal Education and Child Outcomes," *Journal of Human Resources*, 47(4), 873-912, 2012.
- Anger, S., "Intergenerational Transmission of Cognitive and Noncognitive Skills," in J. Ermisch, M. Jäntti and T.M. Smeeding, eds., *From Parents to Children: The Intergenerational Transmission of Advantage*, 393-421, New York, Russell Sage Foundation, 2012.
- Anger, S. and G. Heineck, "Do Smart Parents Raise Smart Children? The Intergenerational Transmission of Cognitive Abilities," *Journal of Population Economics*, 23(3), 1255-1282, 2010.
- Bauer, T.K., M. Fertig and M. Vorell, "Neighborhood Effects and Individual Unemployment," *IZA Discussion Paper No 6040*, IZA Bonn, 2011.
- Bauer, T.K., R. Flake and M.G. Sinning, "Labor Market Effects of Immigration: Evidence from Neighborhood Data," *Review of International Economics*, 21(2), 370-385, 2013.
- Bauer, T.K., T. Kasten and L.H.R. Siemers, "Business Taxation and Wages: Evidence from Individual Panel Data," *IZA Discussion Papers No 6717*, IZA Bonn, 2012.
- Björklund, A., T. Eriksson, M. Jäntti, O. Raum and E. Österbacka, "Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden Compared to the United States," *Journal of Population Economics*, 15(4), 757-772, 2002.
- Björklund, A., K. Hederöf Eriksson and M. Jäntti, "IQ and Family Background: Are Associations Strong or Weak?," *The B.E. Journal of Economic Analysis & Policy - Contributions*, 10(1), Article 2, 2010.
- Björklund, A. and M. Jäntti, "How Important Is Family Background for Labor-Economic Outcomes?," *Labour Economics*, 19(4), 465-474, 2012.
- Björklund, A. and K.G. Salvanes, "Education and Family Background: Mechanisms and Policies," in E.A. Hanushek, S. Machin and L. Woessmann, eds., *Handbook of the Economics of Education* Vol. 3, 201-247, Amsterdam, North-Holland, Elsevier, 2011.
- Carneiro, P., C. Meghir and M. Parey, "Maternal Education, Home Environments, and the Development of Children and Adolescents," *Journal of the European Economic Association*, 11, 123-160, 2013.
- Cunha, F. and J.J. Heckman, "The Technology of Skill Formation," *American Economic Review*, 97(2), 31-47, 2007.
- , "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *The Journal of Human Resources*, 43(4), 738-782, 2008.
- Currie, J., "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 47(1), 87-122, 2009.
- Duncan, G.J., J. Boisjoly and K.M. Harris, "Sibling, Peer, Neighbor, and Schoolmate Correlations as Indicators of the Importance of Context for Adolescent Development," *Demography*, 38(3), 437-447, 2001.
- Duncan, G.J., A. Kalil, S.E. Mayer, R. Tepper and M.R. Payne, "The Apple Does Not Fall Far from the Tree in (Eds.) . Princeton University Press. 23-79.," in S. Bowles, H. Gintis and M. Osborne Groves, eds., *Unequal Chances: Family Background and Economic Success*, 23-79, Princeton, Princeton University Press, 2005.
- Durlauf, S.N., "Neighborhood Effects," in J.V. Henderson and J.-F. Thisse, eds., *Handbook of Regional and Urban Economics* Vol. 4, 2173-2242, Amsterdam, North-Holland, Elsevier, 2004.
- Fortson, J.G. and L. Sanbonmatsu, "Child Health and Neighborhood Conditions: Results from a Randomized Housing Voucher Experiment," *The Journal of Human Resources*, 45(4), 840-864, 2010.
- Galster, G.C., "The Mechanism(S) of Neighbourhood Effects: Theory, Evidence, and Policy Implications," in M. van Ham, D. Manley, N. Bailey, L. Simpson and D. Maclennan, eds., *Neighbourhood Effects Research: New Perspectives*, 23-56, Dordrecht, Heidelberg, London, New York, Springer, 2012.
- Gauck, J., "Inaugural Speech by Federal President Joachim Gauck on March 23, 2012 at the German Bundestag, Berlin," 2012.
- Ginther, D., R. Haveman and B. Wolfe, "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships?," *The Journal of Human Resources*, 35(4), 603, 2000.
- Hawranek, F. and N. Schanne, "Your Very Private Job Agency: Job Referrals Based on Residential Location Networks," *IAB Discussion Paper 1/2014*, IAB Nuremberg, 2014.
- Kling, J.R., J.B. Liebman and L.F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), 83-119, 2007.

- Leventhal, T. and J. Brooks-Gunn, "Moving to Opportunity: An Experimental Study of Neighborhood Effects on Mental Health," *Am J Public Health*, 93(9), 1576-1582, 2003.
- Lindahl, L., "A Comparison of Family and Neighborhood Effects on Grades, Test Scores, Educational Attainment and Income—Evidence from Sweden," *The Journal of Economic Inequality*, 9(2), 207-226, 2011.
- Mazumder, B., "Sibling Similarities and Economic Inequality in the Us," *Journal of Population Economics*, 21(3), 685-701, 2008.
- , "Family and Community Influences on Health and Socioeconomic Status: Sibling Correlations over the Life Course (Contributions)," *The B.E. Journal of Economic Analysis & Policy - Contributions*, 11(3), Article 1, 2011.
- Nicoletti, C. and B. Rabe, "Inequality in Pupils' Test Scores: How Much Do Family, Sibling Type and Neighbourhood Matter?," *Economica*, 80(318), 197-218, 2013.
- Page, Marianne E. and G. Solon, "Correlations between Brothers and Neighboring Boys in Their Adult Earnings: The Importance of Being Urban," *Journal of Labor Economics*, 21(4), 831-855, 2003.
- Raaum, O., K.G. Salvanes and E.O. Sorensen, "The Neighbourhood Is Not What It Used to Be," *The Economic Journal*, 116(508), 200-222, 2006.
- Richter, D., M. Metzger, M. Weinhardt and J. Schupp, "SOEP Scales Manual," *SOEP Survey Papers Series C - Data Documentations No 138*, DIW Berlin, 2013.
- Robert, S.A., "Community-Level Socioeconomic Status Effects on Adult Health," *Journal of Health and Social Behavior*, 39(1), 18-37, 1998.
- Roemer, J.E., *Equality of Opportunity*. Harvard University Press, Cambridge, London, 1998.
- Schnitzlein, D.D., "How Important Is the Family? Evidence from Sibling Correlations in Permanent Earnings in the USA, Germany, and Denmark," *Journal of Population Economics*, 27(1), 69-89, 2014.
- Sieben, I., J. Huinink and P.M. de Graf, "Family Background and Sibling Resemblance in Educational Attainment. Trends in the Former Frg, the Former Gdr, and the Netherlands," *European Sociological Review*, 17(4), 401-430, 2001.
- Solga, H., E. Stern, B. von Rosenblatt, J. Schupp and G.G. Wagner, "The Measurement and Importance of General Reasoning Potentials in Schools and Labor Markets: Pre-Test Report," *DIW Research Notes 10/2005*, DIW Berlin, 2005.
- Solon, G., "Intergenerational Mobility in the Labor Market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics* Vol. 3 Part A, 1761-1800, Amsterdam, New York, and Oxford, Elsevier, North-Holland, 1999.
- Solon, G., M. Corcoran, R. Gordon and D. Laren, "A Longitudinal Analysis of Sibling Correlations in Economic Status," *Journal of Human Resources*, 26(3), 509-534, 1991.
- Solon, G., M.E. Page and G.J. Duncan, "Correlations between Neighboring Children in Their Subsequent Educational Attainment," *Review of Economics and Statistics*, 82(3), 383-392, 2000.
- Tiebout, C.M., "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64(5), 416, 1956.
- vom Berge, P., M. König and S. Seth, "Sample of Integrated Labour Market Biographies (Siab) 1975 - 2010," *FDZ-Datenreport 01/2013EN*, IAB Nuremberg, 2013.
- Wagner, G.G., J.R. Frick and J. Schupp, "The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements," *Journal of Applied Social Science Studies*, 127(1), 139-169, 2007.

Figures and Tables

Figure 1 Sibling and neighbor correlations in youth education and health



Note: the figure shows estimated sibling and neighbor correlations in youth education and health; error bars indicate standard errors; results from REML estimation of a linear multilevel model; standard errors of sibling correlations and neighbor correlations are calculated via the delta method. Neighbor correlations are adjusted for family background characteristics. For details see Tables 4, 5 and 6.

Source: SOEP 2004-2010.

Table 1 Descriptive statistics of education and health outcomes

Variable	Mean	Minimum	Maximum	Standard deviation	Observations
German grade	2.89	1	6	0.82	3708
Math grade	2.97	1	6	1.03	3708
Average grade	2.93	1	6	0.76	3708
Cognitive ability	31.72	3	55	9.33	1715
Mental health	50.01	11.59	72.66	7.84	3088
Physical health	56.25	24.62	68.89	5.12	3088
BMI	21.81	13.05	47.88	3.43	3531
Height	173.21	125	204	9.13	3556

Source: SOEPv29 (2000-2012).

Table 2 Descriptive statistics of parental characteristics

Variable	Mean	Minimum	Maximum	Standard deviation	Observations
Household income	3102.47	10	35000	1756.48	3772
Highest parental education					
No school degree	0.01	0	1	0.12	3772
Secondary school degree	0.19	0	1	0.39	3772
Intermediate school degree	0.31	0	1	0.46	3772
Other degree	0.10	0	1	0.30	3772
University entrance diploma	0.07	0	1	0.26	3772
University graduate	0.31	0	1	0.46	3772

Source: SOEPv29 (2000-2012).

Table 3 Descriptive statistics of neighborhood characteristics

Variable	Mean	Minimum	Maximum	Standard deviation	Observations
Rate of unemployed %	11.35	2.25	36.21	6.23	1722
Rate of debtors %	9.99	2.40	34.18	3.64	1722
Rate of university graduates in the workforce %	7.83	1.41	38.57	4.61	1722
Rate of high school graduates in the workforce %	5.39	1.79	22.52	2.38	1722

Source: IEB 2004-2010; SOEPv29 (2004-2010).

Table 4 Sibling correlations in youth education and health outcomes

	German grade		Math grade		Average grade		Cognitive ability		Mental health		Physical health		BMI	Height		
Sibling correlation	0.17	***	0.23	***	0.22	***	0.46	***	0.31	***	0.21	**	0.33	***	0.42	***
	(0.03)		(0.03)		(0.03)		(0.04)		(0.03)		(0.03)		(0.03)		(0.02)	
Family component	0.11		0.24		0.12		40.35		18.69		5.58		3.89		20.13	
Individual component	0.54		0.82		0.44		46.55		41.54		20.63		7.76		27.68	
Number of families	2664		2664		2664		1303		2249		2249		2524		2538	
Number of individuals	3708		3708		3708		1715		3088		3088		3531		3556	

Note: the table contains estimated sibling correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of sibling correlations are calculated via the delta method.

Source: SOEPv29 2000-2012.

Table 5 Neighbor correlations in youth education and health outcomes

	German grade		Math grade		Average grade		Cognitive ability		Mental health		Physical health		BMI		Height	
<i>A: Raw neighbor correlations</i>																
Neighbor correlation	0.08	***	0.08	***	0.08	***	0.27	***	0.14	***	0.09	***	0.11	***	0.16	***
	(0.02)		(0.02)		(0.02)		(0.04)		(0.02)		(0.02)		(0.02)		(0.02)	
Neighborhood component	0.05		0.08		0.05		23.53		8.22		2.36		1.28		7.67	
Individual component	0.59		0.98		0.52		63.50		51.66		23.84		10.38		40.48	
Number of neighborhoods	1712		1712		1712		1039		1513		1513		1667		1673	
Number of individuals	3708		3708		3708		1715		3088		3088		3531		3556	
<i>B: Neighbor correlations adjusted for household income and highest parental education</i>																
Neighbor correlation	0.06	***	0.08	***	0.07	***	0.22	***	0.13	***	0.09	***	0.10	***	0.15	***
	(0.02)		(0.02)		(0.02)		(0.04)		(0.02)		(0.02)		(0.02)		(0.02)	
Neighborhood component	0.04		0.08		0.04		16.08		7.70		2.15		1.15		7.00	
Individual component	0.58		0.96		0.51		58.43		51.33		23.67		10.29		39.88	
Number of neighborhoods	1712		1712		1712		1039		1513		1513		1667		1673	
Number of individuals	3708		3708		3708		1715		3088		3088		3531		3556	

Note: the table contains estimated neighbor correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of neighbor correlations are calculated via the delta method. Panel A contains raw neighbor correlations, panel B contains neighbor correlations adjusted for household income and highest parental education.

Source: SOEPv29 2000-2012.

Table 6 Adjusted neighbor correlations (sample restricted to youth living in neighborhoods with more than two families)

	German		Math		Average		Cognitive		Mental		Physical		BMI	Height	
	grade		grade		grade		ability		health		health				
Neighbor correlation	0.04	*	0.05	**	0.04	**	0.22	***	0.10	***	0.06	*	0.05	**	0.03
	(0.02)		(0.02)		(0.02)		(0.06)		(0.03)		(0.03)		(0.02)		(0.02)
Neighborhood component	0.02		0.05		0.02		17.44		6.35		1.45		0.51		1.22
Individual component	0.60		1.00		0.54		63.30		55.68		23.94		10.23		40.85
Number of neighborhoods	267		267		267		215		262		262		266		266
Number of individuals	1240		1240		1240		524		1057		1057		1163		1174

Note: the table contains estimated neighbor correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of neighbor correlations are calculated via the delta method. Neighbor correlations adjusted for household income and highest parental education.

Source: SOEPv29 2000-2012.

Table 7 Adjusted neighbor correlations (subsample with neighborhood characteristics)

	German grade	Math grade	Average grade	Cognitive ability	Mental health	Physical health	BMI	Height
<i>Panel A: Neighbor correlations adjusted for household income and highest parental education</i>								
Neighbor correlation	0.09 ** (0.04)	0.09 *** (0.03)	0.09 ** (0.03)	0.21 *** (0.05)	0.19 *** (0.04)	0.09 ** (0.04)	0.13 *** (0.04)	0.14 *** (0.04)
Neighborhood component	0.06	0.10	0.05	15.88	12.07	2.27	1.48	6.50
Individual component	0.57	0.94	0.50	59.19	50.42	23.12	10.17	40.43
Number of neighborhoods	1026	1026	1026	818	934	934	1011	1014
Number of individuals	1697	1697	1697	1263	1475	1475	1659	1672
	German grade	Math grade	Average grade	Cognitive ability	Mental health	Physical health	BMI	Height
<i>Panel B: Neighbor correlations adjusted for household income, highest parental education, neighborhood level of education and economic situation</i>								
Neighbor correlation	0.07 * (0.04)	0.09 *** (0.03)	0.08 ** (0.03)	0.21 *** (0.05)	0.19 *** (0.04)	0.09 ** (0.04)	0.13 *** (0.04)	0.14 *** (0.04)
Neighborhood component	0.04	0.10	0.04	15.40	11.82	2.23	1.47	6.55
Individual component	0.57	0.94	0.50	59.14	50.52	22.97	10.19	40.43
Number of neighborhoods	1026	1026	1026	818	934	934	1011	1014
Number of individuals	1697	1697	1697	1263	1475	1475	1659	1672

Note: the table contains estimated neighbor correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of neighbor correlations are calculated via the delta method. Panel A contains neighbor correlations adjusted for household income and highest parental education, panel B contains neighbor correlations adjusted for household income, highest parental education, unemployment rate, rate of debtors, rate of university graduates, rate of workers with high school diploma, dummy for living in a city with more than 500.000 inhabitants and a dummy for living in East Germany.

Source: SOEPv29 2004-2010; IEB 2004-2010.