Social Mobility in Germany

Majed Dodin (University of Mannheim)
Sebastian Findeisen (University of Konstanz)
Lukas Henkel (European Central Bank)
Dominik Sachs (LMU Munich)
Paul Schüle (LMU Munich and ifo Institute)

Discussion Paper No. 298

November 15, 2021

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Social Mobility in Germany*

Majed Dodin, Sebastian Findeisen, Lukas Henkel,
Dominik Sachs & Paul Schüle

July 15, 2021

Abstract

We characterize intergenerational mobility in Germany using census data on educational attainment and parental income for 526,000 children. Our measure of educational attainment is the A-Level degree, a requirement for access to university. A 10 percentile increase in the parental income rank is associated with a 5.2 percentage point increase in the A-Level share. This parental income gradient has not changed for the birth cohorts of 1980-1996, despite a large-scale policy of expanding upper secondary education. At the regional level, there exists substantial variation in mobility estimates. Place effects, rather than sorting of households, account for most of these differences.

JEL-Codes: I24, J62, R23

Keywords: Intergenerational Mobility, Educational Attainment, Local Labor Markets

*__Dodin:__ University of Mannheim (majed.dodin@gess.uni-mannheim.de); **Findeisen:** University of Konstanz (sebastian.findeisen@uni-konstanz.de); **Henkel:** European Central Bank (lukas.henkel@ecb.europa.eu); **Sachs:** LMU Munich (dominik.sachs@econ.lmu.de); **Schüle:** ifo Munich and LMU Munich (paul.schuele@econ.lmu.de). We thank Klaus Adam, Timm Boenke, Fabian Pfeffer, Winfried Koeniger, Fabian Kosse, Moritz Kuhn, Marc Piopiunik, Uta Schönberg, Jan Stuhler, Ludger Wößmann, Larissa Zierow and seminar participants in Lisbon, Mannheim, Munich, Tübingen and Vienna for helpful comments. We thank Lea Fricke and Victoria Szabo for great research assistance. Financial support by the Joachim Herz Foundation is gratefully acknowledged. The findings and conclusions expressed are solely those of the authors and do not represent the views of the Joachim Herz foundation or of the European Central Bank (ECB).

1 Introduction

Social mobility is an important indicator for both fairness and economic efficiency in a society. Next to violating widely held fairness ideals, low mobility can lead to the misallocation of resources, as talented children from disadvantaged families are impeded from realizing their full potential. Despite its importance, reliable mobility estimates are lacking for many countries. Measuring social mobility is challenging as it requires data that link outcomes of parents to a measure of opportunities for children. Household panel studies may contain this information but are typically too small to deliver estimates for single years or regions with a sufficient degree of statistical confidence (Lee and Solon, 2009; Mazumder, 2018). While in some countries researchers were able to obtain high-quality estimates from administrative tax records (e.g. Chetty et al., 2014), tax data linking children and parents are often unavailable. This is true for Germany as well, where to date no estimates of social mobility across time and space exist that are based on large datasets.

This paper uses census data to document time trends and regional differences in social mobility in Germany at a higher level of detail than previously possible. We measure mobility by the association between the educational attainment of a child and parental income. Our measure of educational attainment, the A-Level (Abitur), is the highest secondary schooling degree in Germany and typically obtained when children still cohabit with their parents. We can therefore estimate mobility statistics using census data, where children can be linked to their parents only if they live in the same household. Our data cover one percent of the German population in every year from 1997 to 2018, providing detailed information on the educational attainment of 526,000 children aged 17 to 21 and the socioeconomic status of their parents.

We present three main findings. First, relative mobility at the national level has remained constant for recent birth cohorts. On average, a 10 percentile increase in the parental income rank was associated with a 5.2 percentage point increase in the probability to obtain an A-Level degree. For the birth cohorts that we observe, 1980-1996, this parental income gradient has not changed even though a large-scale expansion of upper secondary education had been implemented in Germany ("Bildungsexpansion"). This long-term expansion was in parts a policy response to a public debate on social mobility (Dahrendorf, 1966; Hadjar and Becker, 2006) and increased the A-level share from 39% for children born in 1980 to 53% for the 1996 birth cohort. We document that this expansion took place uniformly across the income distribution: the percentage point increase in the A-level share was almost constant across the income distribution. This enhanced the odds ratio for disadvantaged children but left the parental income gradient – that is the absolute difference in the probability to obtain an A-Level – unaffected. The same pattern emerges when estimating mobility trends for population subgroups typically emphasized in social mobility policies, such as children in single parent households or children of parents with low levels of formal education.

Second, we document substantial geographical variation in social mobility across local labor markets. For example, moving up 10 percentile ranks in the national parental income distribution increases the probability to graduate with an A-Level degree by 4.6 percentage points in Hamburg but 6.8 percentage points in Leipzig. While mobility is lower in East Germany, most of the variation originates from within-state differences. This is remarkable, as education policies, which prior literature has suspected to be a key determinant of mobility, vary mainly on the state level in Germany.

Third, we show that household characteristics can explain only a small fraction of these mobility differences. Spatial variation in mobility estimates can either arise due to place effects or due to systematic sorting of households into different local labor markets. Which answer prevails has important implications about the usefulness of place-based policies. For the US, Rothbaum (2016) and Gallagher et al. (2018) suggest that a substantial share of the geographical variation in the intergenerational mobility measures reported in Chetty et al. (2014) can be explained by differences in household characteristics. However, both articles can only indirectly assess the role of household characteristics for mobility estimates, as the administrative tax data used by Chetty et al. (2014) contain little covariates. By contrast, the census data employed in this paper contain rich information on family background. We can therefore directly test whether mobility differences across local labor markets are reduced when controlling for an extensive set of household characteristics. We find that the mobility ranking between local labor markets is largely unchanged when controlling for household characteristics, indicating that rather place-based features account for the geographical variation.

We finally explore which regional characteristics are most predictive for areas of high or low social mobility. To avoid overfitting our data, we use a Random Forest to preselect covariates before estimating a multiple linear model in the spirit of Belloni and Chernozhukov (2013). Drawing on a rich set of regional indicators, we find that local labor market conditions, social characteristics and features of the education system best predict mobility. Regarding the relevance of the education system, the evidence is not clear cut. On the one hand, most of our variation arises within states and the structure and financing of the German education system should theoretically ensure that within-state differences in school quality are modest. On the other hand, we find that the adjusted school drop out rate, a central measure for school quality in Chetty et al. (2014), is the single most predictive indicator for social mobility even in the presence of state fixed effects. Furthermore, labor markets with enhanced possibilities to obtain an A-Level degree from vocational schools are characterized by high mobility.

Apart from providing a comprehensive account of social mobility in Germany, this paper contributes to the literature on the measurement of intergenerational mobility. The current gold standard in measuring intergenerational social mobility is to estimate the expected income rank of children conditional on the income rank of their parents using administrative tax data (Dahl and DeLeire, 2008; Chetty et al., 2014). While administrative tax data has key advantages like a large sample size and a high data quality in the upper half of the income distribution, they come with their own limitations[1] and are often not available to researchers due to tax and data protection laws.

Our approach to measure mobility by the association between the parental income rank and the educational attainment of children after secondary school is an appealing alternative for several reasons. Most importantly, it allows to measure mobility using census data, where children can only be linked to their parents as long as they live in the same household. Census data also improve over standard household surveys in terms of sample size, are representative by construction and – as opposed to administrative tax data – may contain detailed information on the socio-economic situation of households. Moreover, we can obtain mobility statistics for relatively recent cohorts, which is is not possible when estimating income mobility.[2]

[1]Tax return data are not representative, as they capture only the tax-paying part of the population. Another important limitation is the absence of information on household characteristics or other outcomes of interest like education or occupation. See Mazumder (2016; 2018) for a discussion.

[2]A well established literature documents life cycle bias in income mobility estimates (Solon, 1992; Haider and Solon, 2006; Nybohm and Stuhler, 2016). Due to heterogeneity in life cycle earnings profiles, estimates obtained when children are still in their twenties tend to be downward biased. Haider and Solon (2006) suggest to measure the income of children around the age of 40, when this bias is minimized. This implies that even the most recent evidence on income mobility applies to children born 40 years ago. It is then an open question to what extent the insights retrieved for this generation are still relevant for the children born today.

Germany is a country with a rigorous early-age tracking system. This is why the receipt of an A-Level degree is a compelling measure of opportunity in the German context. Only completion of the highest track grants direct access to the tuition-free national university system, opening up the full range of career prospects. As a result, the A-Level wage premium amounts to more than 40%. Besides the economic benefits, having obtained an A-Level is also an important sign of social distinction in the otherwise rather egalitarian German society.

More broadly, a large literature shows that educational attainment has intrinsic value and predicts a wide range of non-pecuniary outcomes (Oreopoulos and Salvanes, 2011; Lochner, 2011). Educational attainment as a measure of opportunity is thus no second best solution in the German context, but, as we believe, a strong and comprehensive indicator for the opportunities of an individual. Beyond Germany, this approach to measure mobility may also prove useful in other countries where the highest secondary school degree plays a similarly important role in shaping future career options.

The remainder of this paper is organized as follows. Section 2 discusses the related literature and summarizes relevant features of the German education system. Section 3 outlines our approach to the measurement of social mobility and how it fits into the German institutional framework. Section 4 describes the data and provides details on the sample construction. In Section 5, we report our results at the national level, while Section 6 shows how mobility has evolved over time. Finally, Section 7 documents geographical variation in mobility. Section 8 concludes.

2 Related Literature and Institutional Background

2.1 Related Literature

Measures of intergenerational social mobility describe the association between opportunities of children and their parents' socioeconomic status. Since economic and social opportunities are difficult to measure, empirical studies of social mobility have generally aimed at the joint distribution of outcomes. The most common approach in the literature focuses on the relationship between income of children and parents, summarized by the intergenerational elasticity of earnings (IGE) (Zimmerman 1992; Solon, 1992; Mazumder, 2005) or more recently by rank-rank correlations (Dahl and DeLeire, 2008; Chetty et al., 2014), which are more robust to non-linearities and measurement issues at the bottom of the income distribution. Both approaches require the researcher to observe child earnings and are therefore only feasible for relatively old birth cohorts. The early canonical studies of intergenerational mobility thus mostly relied on panel data with a long time dimension like the Panel Study of Income Dynamics (PSID) in the US, which allow to observe both parents and children in the age range 30-40.

Given the low sample size of typical household panel data sets like the PSID, mobility estimates are only representative at the national level and it is difficult to detect significant differences between birth cohorts or regions (Mazumder, 2018). A major step forward in terms of data quality has been achieved by Chetty et al. (2014), who use linked administrative tax records to study the association of parent and child incomes. The large sample size of this study has allowed for the first time to obtain reliable estimates of the geographical variation in US intergenerational mobility, opening up the field of research for a better understanding of the causes of mobility (Chetty and Hendren, 2018a; 2018b). This approach was recently replicated also for other countries, including Italy (Acciari et al., 2019), Switzerland (Chuard and Grassi, 2020), Canada (Corak, 2020) and Australia (Deutscher and Mazumder, 2020).

However, linked administrative tax data are not available everywhere. In Germany, it is prohibited by law to link individual tax returns. For that reason, most empirical evidence on social mobility is based on the German Socio Economic Panel (GSOEP), the German counterpart of the PSID. Like the PSID, the GSOEP provides detailed information about child outcomes and parental background but suffers from a low sample size. In the GSOEP it is therefore not possible to document time trends or more fine-grained geographical variation in mobility with a sufficient degree of statistical confidence. Schnitzlein (2016) shows that existing mobility estimates based on the GSOEP are sensitive to small variations in sampling criteria, resulting in a wide range of plausible estimates. It is therefore not surprising that the empirical evidence regarding the level of social mobility in Germany is mixed. Studies that investigate intergenerational income mobility in the GSOEP include Eisenhauer and Pfeiffer (2008), Riphahn and Heineck (2009), Eberharter (2013) and Bratberg et al. (2017). These studies typically find a higher level of mobility than in the US and lower mobility in East than in West Germany, albeit with high statistical uncertainty.

In addition, several studies investigate educational mobility in Germany. Closest to our analysis are Riphahn and Trübswetter (2013) and Klein et al. (2019) who draw on census data to document differences in educational mobility between East and West Germany. While we also use a (more comprehensive) version of this data, we depart from their analysis in considering the joint distribution of child education and parent income. Furthermore, we are able to capture regional differences in much greater detail.

Methodologically, our study is most comparable to Hilger (2015), who for the US also uses the association between the education of children and parental income to leverage the potential of census data for the analysis of intergenerational mobility. As in our data, children are only observed if they co-reside with their parents, posing the risk of a selection bias. The key difference to our approach is that Hilger (2015) measures educational attainment as years of schooling, which typically manifest much later in life than a secondary school degree like the A-Level. Hilger (2015) thus focuses on the age range 26-29, where most children have already left the parental household, and imputes educational attainment for these children. In our setting, this is not necessary as more than 80% of children in the age range 17-21 still live at home and the move-out decision is uncorrelated with parental income. As we argue in the following, the German education system therefore is – compared to the US – much better suited for this type of analysis.[3]

2.2 The German Education System

The German institutional framework is particularly suited to study social mobility through the lens of educational opportunities. We take advantage of the fact that Germany maintains a three-track system of secondary education, where only the highest track grants direct access to the university system. When children finish the four-year[4] elementary school around the age of 10, they are allocated into one of three tracks. Only successful completion of the high track, the Gymnasium (grades 5-12/13), results in the award of an A-Level degree (Abitur). The other two tracks last five (grades 5–9) and six years (grades 5–10) and focus on preparing students for an apprentice-

[3] A small but growing literature on educational mobility in developing countries also relies on children's co-residency with their parents (Asher et al., 2020; Alesina et al., 2021; Muñoz 2021). These studies typically try to minimize selection bias by focusing on simple measures of educational attainment that can already be observed early in life, like having obtained an elementary school degree.

[4] In the states of Berlin and Brandenburg, elementary school lasts six years.

ship. The rigor of the tracking system is mediated by the possibility to switch tracks. In particular, it is common that talented students from the medium track later switch to the general high track or attend a specialized high track. This entails that students "misallocated" after elementary school still have the possibility to correct for this later in their school time.[5]

The responsibility for the education system falls under the jurisdiction of the German states and not under the jurisdiction of the federal government. While the exact implementation of the school system can hence vary across states, these differences do not extend to the A-Level degree. The Standing Conference of State Education Secretaries has the stated goal to ensure a high degree of comparability of educational qualifications across German States and there are no legal differences between the A-Level degrees issued from different states. An A-Level degree grants access to the tuition-free national university system. As schools, universities are state-financed, mostly based on student headcounts. This generates a comparatively large equality in the endowment and quality between different universities.

In consequence, the A-Level degree is by far the most important qualification in the German education system and individuals who obtain it enjoy substantially above-average economic outcomes. Using data on full-time workers aged 30-45, we find an A-Level wage premium of 42% for monthly net income.[6] This estimate mirrors Schmillen and Stüber (2014) who report a 44% A-Level wage premium for total gross lifetime earnings. An A-Level degree is also associated with a lower risk of being unemployed (Hausner et al., 2015) and a higher life expectancy (Gärtner, 2002). Furthermore, it constitutes a beneficial factor for obtaining vocational training in certain white-collar occupations (Klein et al., 2019) and marks an important sign of social distinction in the otherwise rather egalitarian German society. Overall, this illustrates that for children in Germany the A-Level degree is a compelling measure for their social and economic opportunities.

[5] A more detailed overview of the tracking system and track switching in Germany is provided in Biewen and Tapalaga (2017) and Dustmann et al. (2017).

[6] We use the waves 1997-2018 of the German Mikrozensus (described below) and compute the A-Level wage premium by regressing the log of net monthly personal income of full-time workers aged 30-45 on an A-Level dummy, controlling for a full set of age and year fixed effects to implicitly account for job experience. The raw A-Level wage premium is even higher.

3 Measuring Social Mobility

We define two sets of mobility statistics with the aim to distinguish between two distinct mobility concepts: relative and absolute mobility, following the recent literature. Measures of relative mobility compare the outcomes of children from low-income families to those of children from high-income families, whereas measures of absolute mobility indicate the absolute level of opportunity for children from low-income families.

Relative Mobility. Relative mobility is concerned with the difference in opportunities between children from low-income families relative to children from high-income families. In the context of our analysis, this can be summarized by the question:

"How does the probability of obtaining an A-Level degree differ between children from low-income households and children from high-income households?"

To answer this question, we estimate the slope coefficient β in a regression of our binary outcome variable Y_i, indicating whether a child has obtained an A-Level degree, on the parental income rank R_i:

$$Y_i = \alpha + \beta R_i + \varepsilon_i. \tag{1}$$

If the relationship between parental income rank and the probability of obtaining an A-Level degree is well described by a linear relationship (we provide evidence that this is indeed the case later on), the slope coefficient of this regression provides a parsimonious statistic of relative mobility. The probability of obtaining an A-Level degree for a child from a given income percentile r is then given by:

$$E(Y_i|R_i = r) = \alpha + \beta r. \tag{2}$$

and the difference in probabilities between two parental income percentiles r and q, $r > q$, is given by:

$$E(Y_i|R_i = r) - E(Y_i|R_i = q) = \beta(r - q). \tag{3}$$

The slope coefficient thus summarizes the difference in the probability of obtaining an A-Level degree between households from different income percentiles. For better readability, we always multiply β by 100. In what follows, we refer to $\beta \times 100$ as the parental income gradient, our main measure for relative mobility. A high value of the parental income gradient implies a low degree of mobility.

Next to the parental income gradient, we estimate a second statistic of relative mobility, the Q5/Q1 ratio. It is defined as the share of children who obtain an A-Level degree from the top quintile of the parental income distribution divided by the share of children who obtain an A-Level degree from the bottom quintile of the parental income distribution:

$$Q5/Q1 = \frac{E(Y_i|R_i > 80)}{E(Y_i|R_i \leq 20)}. \tag{4}$$

This measure can be interpreted as the relative likelihood of obtaining an A-Level degree between both groups. A ratio of Q5/Q1 = 2 would mean that children from the top quintile of the income distribution are twice as likely to obtain an A-Level degree as children from the bottom quintile of the income distribution. A high value of the Q5/Q1 ratio thus implies a low degree of relative mobility. Conceptually, the Q5/Q1 ratio differs from the parental income gradient in that it is sensitive to the absolute probability of obtaining an A-Level in both quintiles.

Absolute Mobility. Next to relative mobility, one may also be interested in the following question:

"What is the probability of obtaining an A-Level degree for a child from a given (poor) household income level?"

The focus of this question is different from the one before. The question posed here is concerned with the absolute outcomes of disadvantaged children, regardless of the outcomes of children from more advantaged households. We measure absolute mobility by the probability of obtaining an A-Level degree for a child from the bottom quintile of the national income distribution:

$$Q1 = E(Y_i|R_i \leq 20) \tag{5}$$

We refer to this statistic as the Q1 measure. A high value of the Q1 measure implies high absolute mobility.

4 Data

We use the German Microcensus (Mikrozensus, hereafter MZ), an annual representative survey of the German population, administered by the Statistical Office of Germany. The survey is comparable to, but more detailed than, the American Community Survey and constitutes the largest survey program of official statistics in Europe. Each year, 1% of the population living in Germany is randomly selected to participate in the survey. Participation is mandatory for the selected households, which then remain in the survey for four subsequent years. The survey is hence structured as a rotating panel. The MZ covers a wide range of topics, including family status, citizenship, labor market participation, income and educational attainment.[7] In West Germany, the first MZ was administered in 1957, in East Germany in 1991. In our analysis, we use the waves 1997 to 2018, where a consistent definition of the education outcomes of interest is available.

The MZ allows to directly observe family ties within each household. Household members are obliged by law to provide information on every person registered at their respective household. Consequently, we are able to match children to their parents as long as they are still registered at their parents' household. Table 1 displays the share of children living with their parents by age of the child, calculated from our data. Virtually all children younger than 15 still live with their parents. From then on, the share of children co-residing with their parents is decreasing with child age. While 92% of 18 year olds are living with their parents, only 44% of 23 year olds still live at home.

Table 1. Co-Residence Rate by Child Age

Child Age	15	16	17	18	19	20	21	22	23
Share Living with Parents	0.99	0.98	0.97	0.92	0.84	0.72	0.62	0.52	0.44

Notes: This table reports separately by age the fraction of individuals which live in the same household as at least one of their parents in the MZ waves 1997 to 2018.

Measuring Educational Attainment of Children. The primary measure for educational attainment of children in our analysis is a binary variable Y_i that is equal to one if a child has obtained a school degree that qualifies for university entrance or if a child is on track to obtain such a degree, and zero otherwise. Specifically, our child outcome variable is equal to one if (i) a child has obtained a secondary school degree that qual-

[7]For more details on the survey and the sample design see https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Bevoelkerung/mikrozensus-2017.html.

ifies for tertiary education[8] or if (ii) a child is enrolled in the last 2-3 years of a track which leads to such a degree at the successful completion of school.[9] We refer to this outcome as an A-Level degree.

We choose this definition in order to also include younger children in our sample, for which move-out of home is less of an issue than for older children (see Table 1). Furthermore, using this broader definition reduces measurement error in our outcome measure for children aged 18-19. A-Level degrees are usually awarded in the second quarter of the calendar year, while answers to the MZ are collected on a rolling basis. Hence, we have a non-negligible amount of children in our sample that are surveyed before their graduation, but eventually graduate with an A-Level degree in the same year. Back of the envelope calculations suggest that, if we only count children who have already obtained an A-Level degree, we would miss-measure our outcome variable for around 40% of the graduating cohort in the survey year. On the other hand, the share of children failing the final examination in a given year is very low (around 3 percent on average Germany-wide in 2014).[10]

Measuring Parental Income. We measure parental income as the monthly net household income, excluding the monthly net income of all dependent children. This income measure covers all sources of income, including labor income, business profits and social security transfers. Our data contain a continuous measure of household income, which we use as our main income measure. This income measure is not asked directly in the survey, but imputed by the Statistical Office: Respondents in the MZ report both their personal and their household income in 24 predefined bins. The Statistical Office then transforms the personal binned income into a continuous variable, essentially assuming that individuals are equally distributed within each bin. In a second step, these values are summed up to a continuous measure of household income. We thus measure parental income with some imprecision. This measurement error is independent from the educational outcomes of children but could theoretically lead to a small attenuation bias. We would then slightly underestimate the influence of parental background (overestimate mobility) in all of our measures. However, Nybom and Stuhler

[8]This includes *Allgemeine Hochschulreife (Abitur)*, *Fachgebundene Hochschulreife* and *Fachhochschulreife*.

[9]This definition subsumes all students on *Allgemeinbildende Schulen* enrolled in the *Gymnasiale Oberstufe* as well as students from specialized tracks like *Berufliches Gymnasium* or *Fachoberschule* which award an A-Level degree.

[10]For an overview of the share of children failing the final examination see https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html

(2017) show that rank-based mobility measures like the ones employed in this paper are very robust to attenuation bias.

To account for differences in need and standard of living by household composition we scale all household incomes by the modified OECD equivalence scale. Appendix A shows that the choice of a scaling factor is not influential for our results. Finally, following Dahl and DeLeire (2008) and Chetty et al. (2014), we assign every household its percentile rank in the income distribution relative to all other households with children in the same age range within our sample, separately for every year. We refer to this variable as the parental income rank R_i.

Table 2. Summary Statistics by Age of Children

Child Age	Share Female	Mean Parental Income (Equiv.)	Parental Income Rank	Share Parents with A-Level
13	0.49	1153	47	0.35
14	0.49	1127	45	0.34
15	0.49	1167	47	0.34
16	0.49	1161	46	0.33
17	0.49	1244	50	0.33
18	0.48	1245	50	0.32
19	0.47	1245	50	0.32
20	0.44	1239	50	0.31
21	0.42	1243	50	0.31
22	0.41	1162	46	0.31
23	0.39	1174	46	0.30

Notes: This table reports separately by age summary statistics for children living with their parents in the MZ waves 1997 to 2018. As in our baseline estimates, parental income is equivalized using the modified OECD scale. A summary statistic that systematically changes with age is correlated with the move-out of children, while a statistic that does not change is uncorrelated.

Sample Definition and Summary Statistics. In our primary sample, we restrict our analysis to children aged 17 to 21, living in the same household as at least one parent. The age cut-offs of 17 and 21 are chosen to balance the following trade-off: For older children, our outcome variable is measured more precisely, i.e. we do not need to rely on enrollment in the high track but are more likely to observe the completed degree. However, the older the children, the higher the fraction that already moved out of the parental household. We therefore choose the higher cut-off at 21 because move-out increases strongly after this age. The lower cut-off is chosen as children enrolled in an

A-Level track are typically at least 17 years old. Despite this conservative age range, we could under- or overestimate mobility with our measurement approach if the decision to move-out were systematically correlated with parental income and the educational attainment of children. Table 2 suggests that this is not the case by showing how key time-constant characteristics of the children in our sample change with age. If move-out would be random, we should not see systematic changes in these statistics for older children. Unsurprisingly, move-out is not random and stratified along various characteristics such as gender. However, there is no significant change in the mean parental income rank from age 17, where 97% of children still live at home, to age 21, where only 62% of children still live with their parents. This suggests that sample selection is not a major concern for our analysis. More importantly, we demonstrate in the next section that choosing an alternative age range barely changes our results.

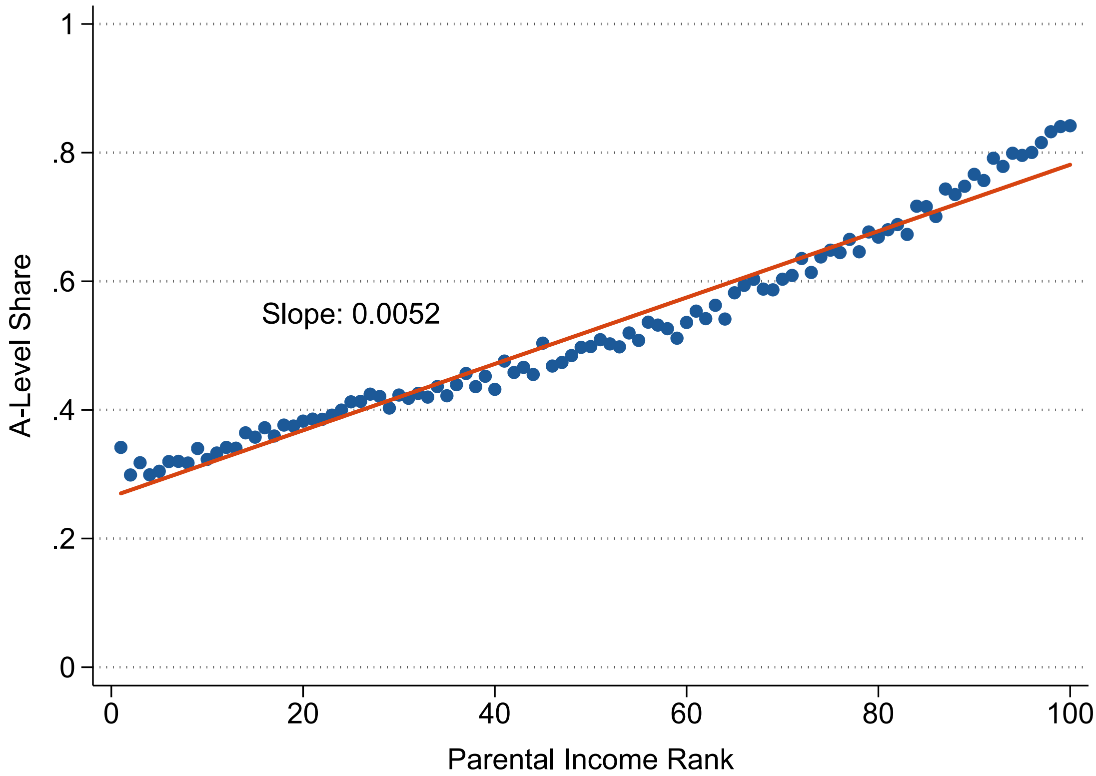
5 National Estimates

In this section, we characterize social mobility at the national level. Figure 1 shows how the share of children with an A-Level degree varies with the position in the parental income distribution.[11] The graph reveals an essentially linear relationship. As a consequence, the empirical conditional expectation function of obtaining an A-Level is well approximated by the slope coefficient as estimated in Equation 1. We find a slope coefficient of $\beta = 0.0052$, which translates into a parental income gradient of $\beta \times 100 = 0.52$. This means that a 10 percentile increase in the parental income rank is associated with an increase in the probability of obtaining an A-Level by 5.2 percentage points.

As outlined in Section 3, we also compute quintile-based measures of intergenerational mobility. We find Q1 = 0.34, meaning that in the bottom quintile of the parental income distribution 34% of children eventually attain an A-Level degree. This compares to 52% on average and 76% in the top quintile of the parental income distribution (Q5 = 0.76). Accordingly, the Q5/Q1 ratio, our second measure of relative mobility, is 0.76/0.34 = 2.24: A child in the top quintile of the parental income distribution is more than twice as likely to graduate with an A-Level than a child in the bottom quintile.

[11] For the national estimates, we pool our data in the period 2011-2018 to achieve a high sample size for the most recent years. As shown later in this paper, the year range chosen here is not influential as the parental income gradient remained remarkably stable for the time of our analysis.

Figure 1. Social Mobility at the National Level



Notes: This figure shows a non-parametric binned scatter plot of the fraction of children aged 17-21 that are either enrolled in the last two/three years of an A-Level track or already attained an A-Level degree by percentile rank of their parents in the national income distribution in the period 2011-2018. The parental income rank is computed within each year among all households with children aged 17-21. The OLS slope of 0.0052 reported in the figure is estimated using the underlying micro data.

Table 3 shows that these estimates are robust to choosing an alternative age range: If we only select children aged 17-19 into our sample, the estimates of the unconditional A-Level share, the Q1 and Q5 measures and the gradient change only in the third decimal place. This shows that the move-out of many 20 and 21 year olds (compare Table 1) does not introduce a significant bias. Likewise, the estimates for children aged 19-21 differ only marginally. Once we select children aged 16-22, we underestimate the A-Level share as many 16 year olds – mostly from high-income households – are still in the school year prior to the A-Level track and are thus misclassified by our measure. Yet, also in this sample we estimate an only slightly lower parental income gradient of $\beta \times 100 = 0.49$.

Do these estimates depict Germany as a country of high or low mobility? A cross-country comparison of our results is not straightforward, as the institutional setting exploited in this paper is quite unique. For the US, we are aware of two studies which

Table 3. Robustness of National Estimates

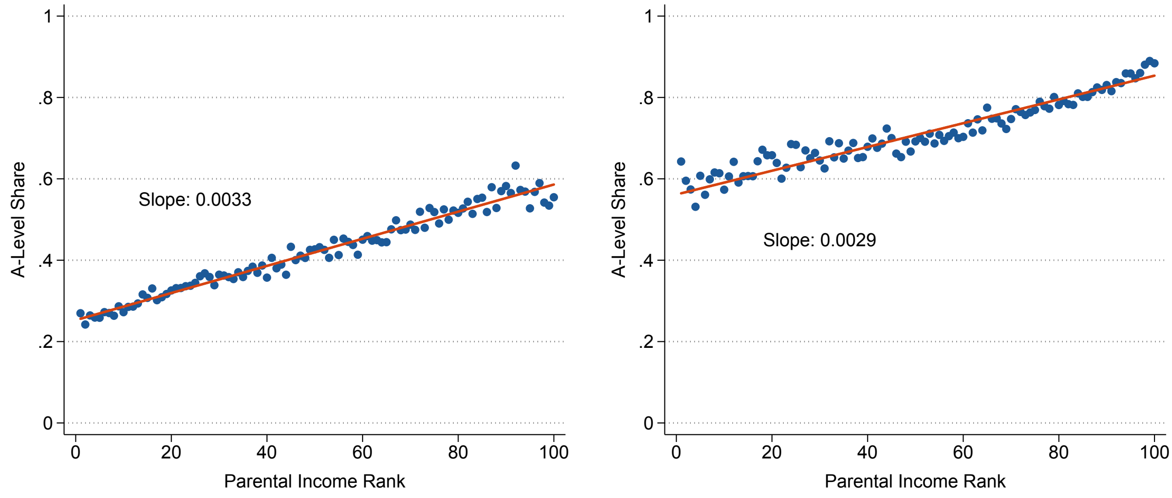
Age	A-Level Share	Q1	Q5	Q5/Q1	Gradient	N
17-21	0.52	0.34	0.76	2.24	0.52	224,017
16-22	0.48	0.31	0.71	2.30	0.49	307,050
17-19	0.52	0.34	0.76	2.26	0.52	147,035
19-21	0.53	0.35	0.76	2.20	0.51	122,345

Notes: This table displays mobility statistics at the national level for MZ waves 2011-2018. The results in the first row (age range 17-21) are our baseline estimates. The bottom three rows show how these statistics change once we select a different age range for the children in our sample.

investigate similar relationships. Hilger (2015) uses data from the Census 2000 and finds for children aged 19-21 that a 10 percentile increase in parental income rank is associated with an increase in the share of children who attend college by 3.6 percentage points. A much higher number is reported in Chetty et al. (2014), who find an association between parental income rank and college enrollment of 6.7 percentage points. Both estimates compare to the 5.2 percentage points for the share of children obtaining the possibility to attend college in our setting. Under the assumption that the likelihood to attend college conditional on having obtained an A-Level degree is weakly increasing in the parental income rank, we would expect an association between parental income rank and college enrollment of at least 5.2 percentage points in Germany. Therefore, the mobility estimates documented for Germany in this article fall within the same range as recent evidence for the US, where researchers have for a long time stressed the difficulties of children from low-income households to access college.

Subgroup Estimates. Next, we analyze how mobility varies between different population subgroups. We continue to rank parents according to their position in the aggregate income distribution. Figure 2 displays the parental income gradient separately for children from households where no parent has an A-Level degree (Panel a) and for children from households where at least one parent has an A-Level degree (Panel b). As expected, children with low parental education have a much lower baseline probability to obtain an A-Level. Remarkably, this difference is close to constant across the whole income distribution and the parental income gradients in both groups are very similar: $\beta \times 100 = 0.33$ for children without parental A-Level degree and $\beta \times 100 = 0.29$ for

Figure 2. Differences by Parental Education



(a) No Parent with A-Level (b) At Least One Parent with A-Level

Notes: This figure shows a binned scatter plot of the fraction of children aged 17-21 where none of the parents has obtained an A-Level degree (Panel a) or where at least one of the parents has obtained an A-Level degree (Panel b) that are either enrolled in the last two/three years of an A-Level track or already completed an A-Level degree by percentile rank of their parents in the national income distribution in the period 2011-2018. The parental income rank is computed within each year among all households with children aged 17-21. The OLS slopes reported in the figure are estimated using the underlying micro data.

children with parental A-Level degree. Furthermore, the probability of obtaining an A-Level still rises linearly in the parental income rank in the two groups defined by our conditioning variable.

Next to parental education, we also explore differences between other subgroups of the German population, summarized in Table 4. The baseline probability to obtain an A-Level degree is lower for males, migrants[12] and children late in the birth order, as well as for children where only one parent is present or the parents are not married. The parental income gradient differs mainly by migration background and the parental marital status. For children where parents are not married, the association between parental income and the probability to obtain an A-Level is particularly strong. The opposite holds true for migrants, where parental income is comparatively less important. Interestingly, the Q1 measure is higher for migrants than for natives. In addition, relative mobility is lower in East Germany than in West Germany. We investigate this regional pattern in more detail in Section 7.

[12]Migrants are defined as all individuals who immigrated to Germany after 1949, as well as all foreigners born in Germany and all individuals born in Germany with at least one parent who immigrated after 1949 or was born in Germany as a foreigner.

Table 4. Social Mobility for Subgroups

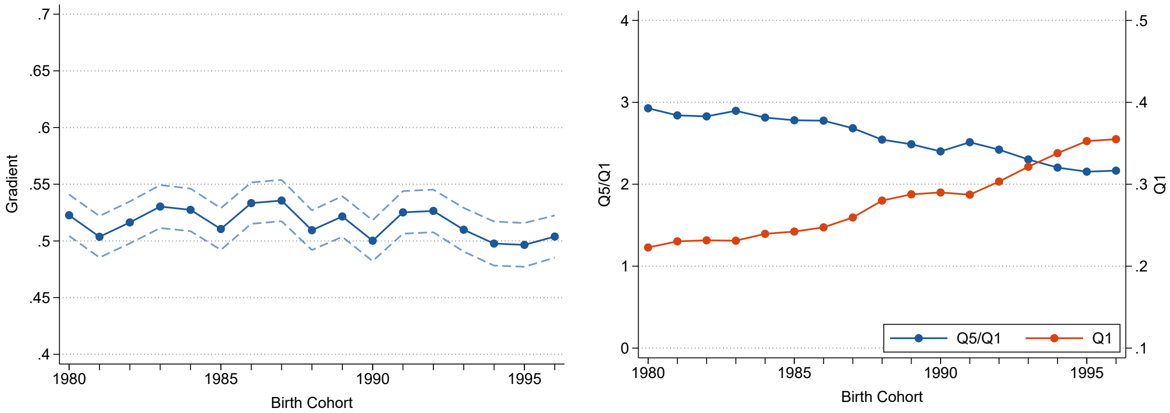
		A-Level Share	Q1	Q5	Q5/Q1	Gradient	N
Gender	Male	0.47	0.29	0.72	2.49	0.53	119,975
	Female	0.58	0.40	0.81	2.02	0.50	104,042
Parental Education	No A-Level	0.39	0.29	0.55	1.93	0.33	141,148
	A-Level	0.75	0.61	0.84	1.37	0.29	82,869
Migration Background	Native	0.54	0.32	0.76	2.35	0.55	163,018
	Migrant	0.48	0.36	0.75	2.11	0.47	60,908
Parenting Status	Single Parent	0.47	0.34	0.72	2.12	0.50	49,397
	Two Parents	0.54	0.34	0.76	2.26	0.55	174,620
Parents Married	Not Married	0.47	0.33	0.69	2.11	0.46	51,018
	Married	0.54	0.35	0.77	2.22	0.54	172,999
Region	East Germany	0.51	0.31	0.80	2.60	0.60	28,872
	West Germany	0.52	0.35	0.76	2.19	0.51	195,145
Siblings	Yes	0.52	0.35	0.79	2.29	0.55	151,793
	No	0.52	0.32	0.72	2.26	0.49	72,224
Birth Order	First Child	0.53	0.34	0.76	2.21	0.51	160,884
	Second Child	0.51	0.34	0.77	2.27	0.52	54,978
	Later Child	0.45	0.32	0.78	2.47	0.57	8,155

Notes: This table shows estimates of relative and absolute mobility for selected subgroups in the MZ waves 2011-2018. Migration background subsumes all individuals who immigrated to Germany after 1949, as well as all foreigners born in Germany and all individuals born in Germany with at least one parent who immigrated after 1949 or was born in Germany as a foreigner. Parental marital status indicates if both parents are married.

6 Time Trends

How has social mobility in Germany evolved over time? To answer this question, we focus on a sample of 526,000 children born between 1980-1996.[13] Compared to most other studies, we therefore consider relatively recent birth cohorts. Figure 3 depicts the evolution of our mobility statistics for these birth cohorts. As illustrated in Panel (a), the parental income gradient remained remarkably stable over all cohorts under

Figure 3. Mobility Statistics by Cohort



(a) Parental Income Gradient

(b) Q5/Q1 Ratio and Q1 Measure

Notes: This figure displays time trends for different statistics of social mobility for children aged 17-21 over birth cohorts 1980-1996. Panel (a) displays the the parental income gradient, computed as $\gamma_t \times 100$, where γ_t is estimated by OLS in the following equation $Y_{i,t} = \alpha + \beta_t C_t + \gamma_t C_t \times R_i + \varepsilon_{i,t}$, with C_t denoting a cohort and $C_t \times R_i$ the interaction between cohort dummies and parental income rank. The dashed lines show 95% confidence bands. Panel (b) depicts the Q5/Q1 ratio and the Q1 measure.

consideration. At the same time, the Q5/Q1 ratio decreased moderately, while the Q1 measure increased (Figure 3, Panel b).

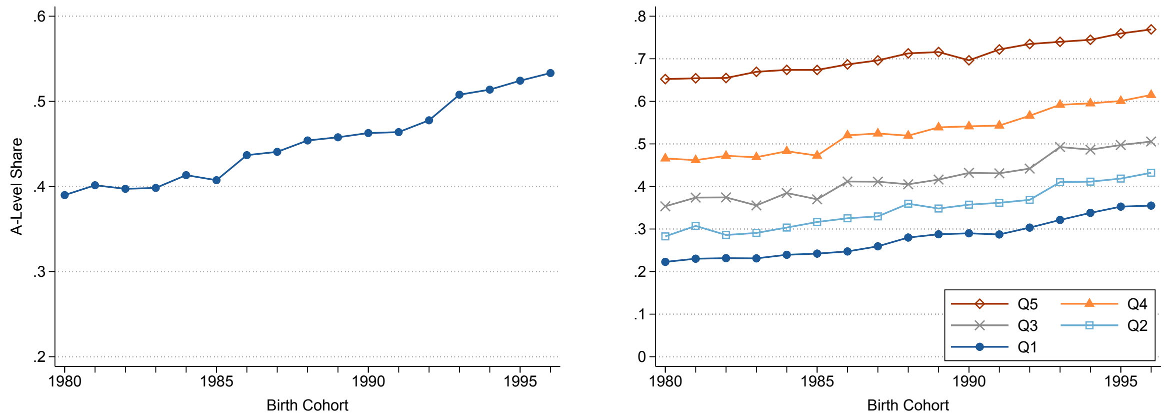
These statistics reflect two salient features of the development of upper secondary education in Germany. First, children born during these years faced a marked increase in the share of awarded A-Levels. As illustrated in Panel (a) of Figure 4, the share of children with an A-Level degree rose from 39% for children born in 1980 to 53% for children born in 1996, implying an increase of about one third. This increase was part of the "Bildungsexpansion", a large-scale policy of expanding higher education in Germany that, starting in the early 1970s, increased the A-Level share from 20% to around 50%.[14] The reform was responding to a public debate on social mobility and the increasing importance of education for economic growth (Dahrendorf, 1966; Picht, 1964).

Second, as illustrated in Panel (b), the absolute increase in the share of awarded A-Level degrees was almost identical in each quintile of the parental income distribution.

[13]We restrict our attention to these cohorts to rule out that the trends reported in this section may be driven by cohort differences in the distribution of age at measurement. For the considered cohorts, the share of 17, 18-, 19-, 20- and 21-year-olds in our data is constant.

[14]The "Bildungsexpansion" did not only increase upper secondary but also tertiary education and we find no evidence that the share of A-Level graduates taking up university studies changed in the recent past. In the years 2002-2015, where most of our birth cohorts graduate, it fluctuated around 70%. See https://www.datenportal.bmbf.de/portal/de/Tabelle-2.5.74.html.

Figure 4. A-Level Share by Cohort



(a) A-Level Share

(b) A-Level Share by Quintile of the Parental Income Distribution

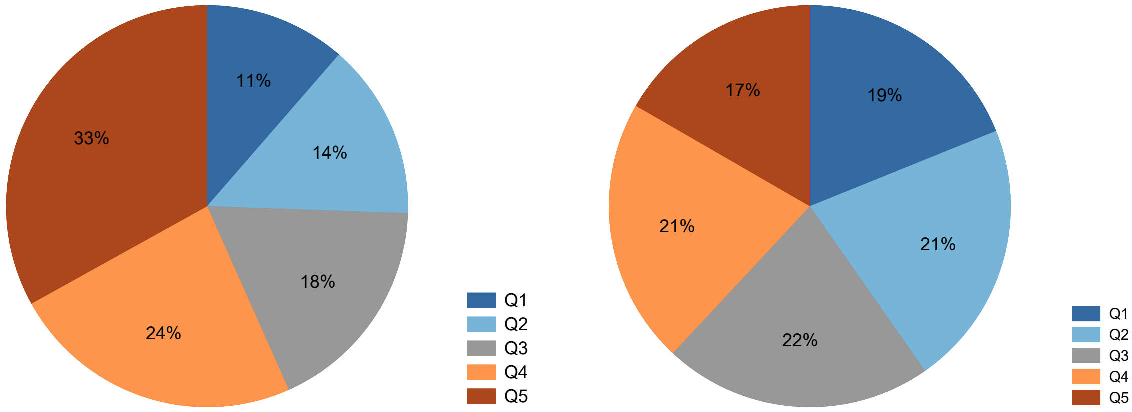
Notes: This figure shows the development of the A-Level share, defined as the fraction of children aged 17-21 that are either enrolled in the last two/three years of an A-Level track or already completed an A-Level degree, for birth cohorts 1980-1996. Panel (a) displays the development for all children, whereas Panel (b) reports the A-Level share separately for each quintile of the parental income distribution.

Although the share of children in the top and bottom quintile increased somewhat below average (12% and 13% respectively compared to 14% on average), the expansion of the A-Level is best described as uniform across quintiles.

The same patterns emerge when focusing on population subgroups. In Appendix B, we plot time trends for the A-Level share and the parental income gradient separately by gender, region, parental education and marital status, the number of siblings and the number of parents in the household. The A-Level share rose equally strong in all subgroups (Figure B.1). Likewise, we do not observe diverging trends in the parental income gradient within subgroups (Figure B.2). We conclude that all parts of the population participated in the expansion of the A-Level share and that the stability of relative mobility over time does not mask important counteracting developments between subgroups.

To illustrate how a roll-out of higher secondary education affects our mobility measures, we conduct a simple thought experiment. Assume for concreteness that the increase in the A-Level share was achieved by opening one additional class in an A-Level track. This gives additional (marginal) children the opportunity to obtain an A-Level. If only one of those marginal children belongs to a family from the bottom quintile of the parental income distribution, the Q1 measure increases. Thus, absolute mobility as we measure it tends to mechanically rise in the A-Level share. The reactions of the two relative measures of mobility depend on the parental background of *all* children entering the marginal class.

Figure 5. A-Level Share Composition by Parental Income



(a) Average Class in 1980 (b) Average Marginal Class 1980-1996

Notes: This figure decomposes the A-Level share, defined as the fraction of children aged 17-21 that are either enrolled in the last two/three years of an A-Level track or already completed an A-Level degree, by quintile of the parental income distribution. Panel (a) shows the average class composition 1980, capturing children which are (or recently have completed) a class on an A-Level track. Panel (b) shows the average "marginal" class composition in the period 1980-1996, capturing the additional children which account for the increase of the A-Level share in this period.

Figure 5 depicts our thought experiment. Panel (a) illustrates the distribution of household income for an A-level class of the 1980 cohort: 33% of the students are from Q5, whereas only 11% are from Q1. Under the assumption that this composition had been the same for the 1996 cohort in the absence of the educational expansion, we can infer what the distribution of household incomes looks like for a 'marginal A-level class' that resulted from the expansionary policy. As shown in Panel (b), selection into the marginal class happened independently of parental income.[15] This explains the different trends of the parental income gradient and the Q5/Q1 ratio: While the uniform expansion altered the absolute probabilities of obtaining an A-Level (and therefore the Q5/Q1 ratio) it did not affect the expected difference in ranks between a child with and without an A-level degree, and therefore did not alter the parental income gradient.

When interpreting our two statistics of relative mobility, is important to bear in mind that the decline in the Q5/Q1 ratio is mainly driven by the diverging baseline

[15] The numbers result from the absolute increase in each quintile (see Figure 4, Panel b) normalized by the total increase in the A-Level share between two birth cohorts.

probabilities in the respective quintiles. Another angle to approach the same development would be to compute the probability *not to obtain* an A-Level for children in both quintiles. For the birth cohort 1980, children in Q1 where 2.2 times more likely not to obtain an A-Level degree as children in Q5. For children born in 1996, this inverse odds ratio has increased to 2.8, meaning that the relative gap in not obtaining an A-Level had actually widened. The parental income gradient in contrast summarizes the absolute A-Level gap between high and low income children and is therefore insensitive to the chosen reference point.

In summary, we conclude that the expansion of upper secondary education was uniform across quintiles, lead to an attenuation of the odds ratio but did not change the parental income gradient. The question whether the documented developments are good or bad news with respect to the efficiency of the German education system depend on the unknown joint distribution of ability and parental income as well as the effective selection mechanism generating our data. If selection was based solely on an ability index, the observed development implies equality of the conditional ability densities for each quintile at each marginal level of ability. In other words, for each marginal level of ability, there were as many not yet enrolled children from the top quintile as from any other quintile. We do not take a stance here on whether this is plausible but emphasize that efficiency considerations require knowledge of quantities that are difficult to estimate.

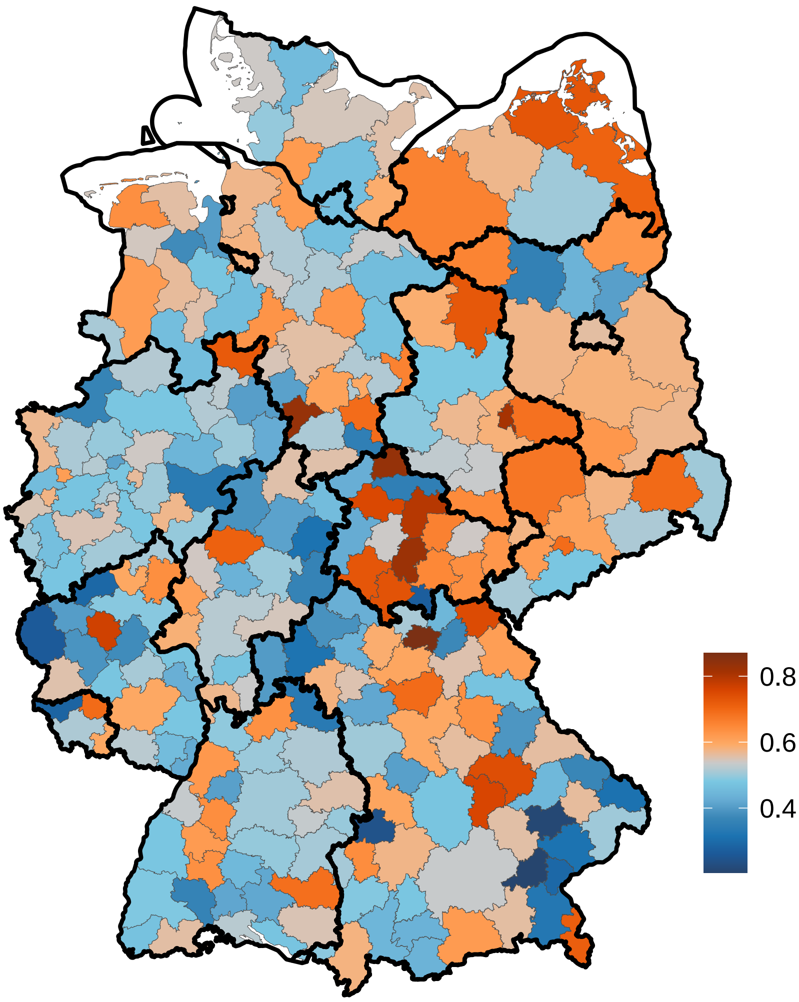
7 Geographical Variation

We now turn to a detailed characterization of regional differences in social mobility in Germany. The unit of analysis is the local labor market (LLM). The 258 LLMs in Germany present an aggregation of counties based on commuting flows, comparable to the definition of commuting zones in the US. With the exceptions of the five LLMs Bremen, Bremerhaven, Hamburg, Mannheim and Ulm, all are strict subsets of states. We assign households to the LLM of their current place of main residence as reported in our data.[16] The median number of children in our sample (observations) per LLM is 552 (mean: 895). The lowest number of observations across all LLMs is 100 (LLM Sonneberg) and the largest number of observations is 8159 (LLM Stuttgart). For each LLM,

[16]Conceptually, this is very similar to the approach in Chetty et al. (2014), who use the place of residence of a child at the age of 15 as their main geographical indicator.

we separately compute the parental income gradient, the Q1 and Q5/Q1 measures and the unconditional A-Level share. Again, we continue to rank parents according to their position in the national income distribution.

Figure 6. Parental Income Gradient by Local Labor Market



Notes: This figure presents a heat map of the parental income gradient by LLM. Children are assigned to LLMs according to their current place of residence. The estimates are based on children aged 17-21 in the years 2011-2018 for which we have non-missing information on educational attainment and parental income. The parental income gradient is obtained as the slope coefficient of a regression of the A-Level dummy on a constant and the parental income rank, multiplied by 100.

Differences in Relative Mobility. Figure 6 presents a heat map of relative mobility.[17] Blue areas represent areas of high mobility, whereas red areas indicate a high gradient

[17]We focus on the parental income gradient. The corresponding heatmap for the Q5/Q1 ratio is displayed in Figure C.1 in the appendix.

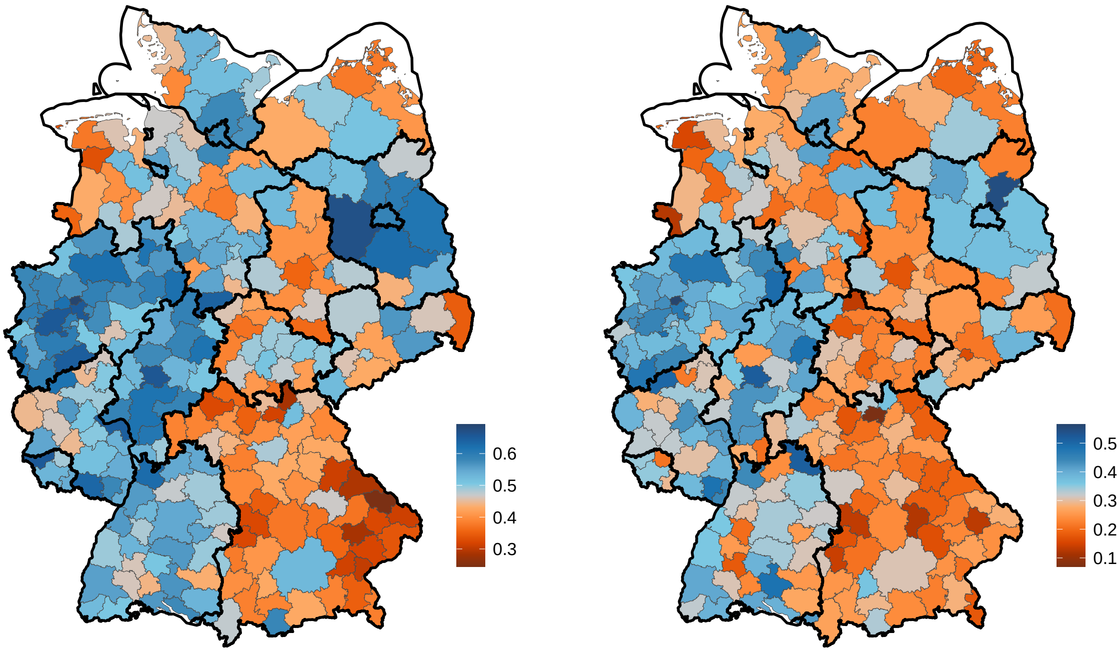
and thus less mobile regions. While LLMs in the East exhibit lower mobility on average, clusters of high and low mobility are spread out across all of Germany. In some rural LLMs the parental income gradient ranges below 0.3, whereas in the least mobile areas the gradient exceeds 0.8. In these regions, a child born at the top of the income distribution has a 80 percentage point higher chance to obtain an A-Level degree than a child born at the bottom of the distribution. The LLMs with the highest gradient (Lichtenfels) and the lowest gradient (Mühldorf) are both located in Bavaria, highlighting the existence of substantial variation in mobility even within states. Indeed, we find that only 12% of the variation across LLMs can be explained by state level differences.[18]

When comparing relative mobility across different regions of Germany, one may be concerned that computing the income ranks of parents in the national income distribution could yield undesirable results in LLMs where the local income distribution covers only a subset of the national one. For example, if in a rural LLM all parental income ranks fell between the 30th and 50th percentile of the national income distribution, a moderate A-Level gap between the children from the poorest and richest families of 15 percent would translate into a very high parental income gradient of 0.75. To address this concern, we compare in Appendix D the distribution of the parental income gradients computed at the national level to the distribution of the parental income gradients computed at the state level. Reassuringly, both estimates are very similar results and the geography of social mobility in Germany remains virtually unchanged when computing the parental income ranks in the state specific income distribution.

Differences in Absolute Mobility. These large within-state differences are remarkable, as the unconditional A-Level share is clearly clustered at the state level. As illustrated in Figure 7, Panel (a), the share of children who eventually obtain an A-Level is systematically higher in Northrine-Westfalia, Hessia and the city states Berlin, Bremen and Hamburg. Similar patterns are also visible for our measure of absolute mobility in Panel (b). Again, red areas indicate regions of low and blue values regions of high mobility. Absolute mobility is lower in Bavaria and higher in Northrine-Westfalia compared to the German average. Just visually, it hence becomes evident that the uncondi-

[18]This calculation is based on a regression of the slope in each LLM on a set of state fixed effects. This analysis uses only LLMs that are strict subsets of states, i.e. the 5 LLMs that cross state boundaries are excluded from this analysis. The R-squared of this regression is $R^2 = 0.1218$, meaning that roughly 12% of the variation arises from state-level differences.

Figure 7. A-Level Share and Q1 Measure by Local Labor Market



(a) A-Level Share (b) Q1 Measure

Notes: This figure presents heat maps of the A-Level share (Panel a) and the Q1 measure (Panel b) by LLM. Children are assigned to LLMs according to their current residence. The estimates are based on children aged 17-21 in the years 2011-2018 for which we have non-missing information on educational attainment and parental income. The A-Level share is defined as the fraction of children aged 17-21 that are either enrolled in the last two/three years of an A-Level track or already completed an A-Level degree. The Q1 measure reports this same share for children in the bottom 20% of the parental income distribution.

tional A-Level share is closely linked to absolute but not to relative mobility. Overall, we also find substantial variation in absolute mobility. In some regions, less than 15% of children from the bottom quintile of the national income distribution obtain an A-Level degree, whereas in other regions this number exceeds 50%. We find that 42% of the variation in the Q1 measure and 57% of the variation in the A-Level share can be attributed to state level differences.

Mobility in the Largest Local Labor Markets. Table 5 presents estimates of absolute and relative mobility for the 15 largest LLMs in Germany. Labor markets are ranked by parental income gradient, which varies between 0.46 in Hamburg and 0.68 in Leipzig. The probability to obtain an A-Level is thus substantially less dependent on parental income for children raised in Hamburg compared to children raised in Leipzig. In

Table 5. Mobility in the 15 Largest Local Labor Markets

Rank	Name	Population	Gradient	Q5/Q1	Q1	A-Level Share
1	Hamburg	2911254	0.46	1.93	0.41	0.58
2	Düsseldorf	1552097	0.47	1.89	0.45	0.65
3	Münster	810377	0.48	1.78	0.47	0.62
4	Gelsenkirchen	1150983	0.50	2.01	0.40	0.57
5	Stuttgart	2522246	0.50	2.23	0.34	0.55
6	Bonn	924546	0.51	1.95	0.44	0.65
7	Duisburg	1170198	0.51	2.02	0.42	0.58
8	Frankfurt/Main	2274907	0.52	1.96	0.43	0.62
9	München	2764232	0.53	2.31	0.31	0.53
10	Dortmund	1159719	0.54	2.14	0.40	0.59
11	Köln	1830947	0.55	2.24	0.38	0.60
12	Hannover	1152675	0.56	2.51	0.30	0.53
13	Berlin	3613495	0.56	2.19	0.39	0.59
14	Nürnberg	1094554	0.60	3.01	0.23	0.43
15	Leipzig	1037782	0.68	3.14	0.25	0.48

Notes: This table reports estimates of absolute and relative mobility for the 15 largest LLMs in Germany (measured by total population in 2017). LLMs are sorted in descending order by the relative mobility rank, as measured by the parental income gradient. A higher rank thus indicates higher relative mobility.

terms of absolute mobility, the variation across LLMs is even larger: the probability to obtain an A-Level for a child in the bottom 20% of the income distribution is nearly twice as high in Hamburg (41%) as in Leipzig (25%) or Nürnberg (23%). Table 5 also illustrates that a high degree of absolute mobility does not necessarily imply a high degree of relative mobility. Stuttgart for example, which is ranked 5th in terms of relative mobility would only be ranked as number 11 when considering absolute mobility. The opposite is true for Berlin, which improves from rank 13 to 9 when moving from relative to absolute mobility. Appendix E complements this evidence by showing mobility estimates for all German states. Children born in the state of Hamburg face the highest degree of social mobility. With the exception of Bremen, the least mobile states are all located in East Germany.

Correlation of Mobility Measures. Table 6 compares how these different mobility statistics relate to each other. As expected, the A-Level share is closely correlated with the Q1 measure: in LLMs with a high A-Level share, children in the bottom quintile have a comparatively high likelihood to complete an A-Level degree. The Q5/Q1 ra-

Table 6. Correlation between Mobility Statistics

Measure	Corr.	A-Level	Q1	Q5/Q1	Gradient
A-Level	ρ	1	-	-	-
	r	1	-	-	-
Q1	ρ	0.75	1	-	-
	r	0.77	1	-	-
Q5/Q1	ρ	-0.39	-0.73	1	-
	r	-0.45	-0.84	1	-
Gradient	ρ	-0.03	-0.47	0.66	1
	r	-0.09	-0.49	0.77	1

Notes: This table reports the correlation between estimates of different measures of social mobility across LLMs in Germany. ρ denotes the Pearson correlation coefficient of two measures across LLMs, r denotes the Spearman rank correlation coefficient.

tio in turn is negatively correlated with both the A-Level share and the Q1 measure. In contrast, there exists no systematic association between the A-Level share and the parental income gradient, because the gradient is not sensitive to the baseline probability of obtaining an A-Level degree. For the same reason, the correlation between the parental income gradient and the Q5/Q1 ratio is – although clearly positive – not close to one. The correlation between the parental income gradient and the Q1 measure even ranges below 0.5, which demonstrates that a high level of absolute mobility in a given LLM does not always imply a high level of relative mobility. Altogether, these cross-sectional correlations mirror our considerations regarding the time trends in the previous section. While the A-Level share is strongly linked to the quantile measures, it's correlation with the parental income gradient is negligible.

7.1 Places versus Sorting

What drives these large mobility differences across local labor markets? One explanation are place effects. An active literature argues that places can shape economic outcomes and that place-based policies can help to improve local conditions (Kline and Moretti, 2014; Neumark and Simpson, 2015). However, regional differences can also arise if household systematically sort into LLMs. If, for example, (i) parental education differs across areas because households sort into different LLMs based on their education status and (ii) parental income and education are correlated, spatial mobility differences would arise even in absence of place effects. For the US, Rothbaum

(2016) and Gallagher et al. (2018) suggest that a substantial share of the geographical variation in the intergenerational mobility measures reported in Chetty et al. (2014) can be explained by differences in household characteristics across commuting zones. However, this is challenging to verify with the data used by Chetty et al. (2014), as administrative tax data provide only limited information on household characteristics.

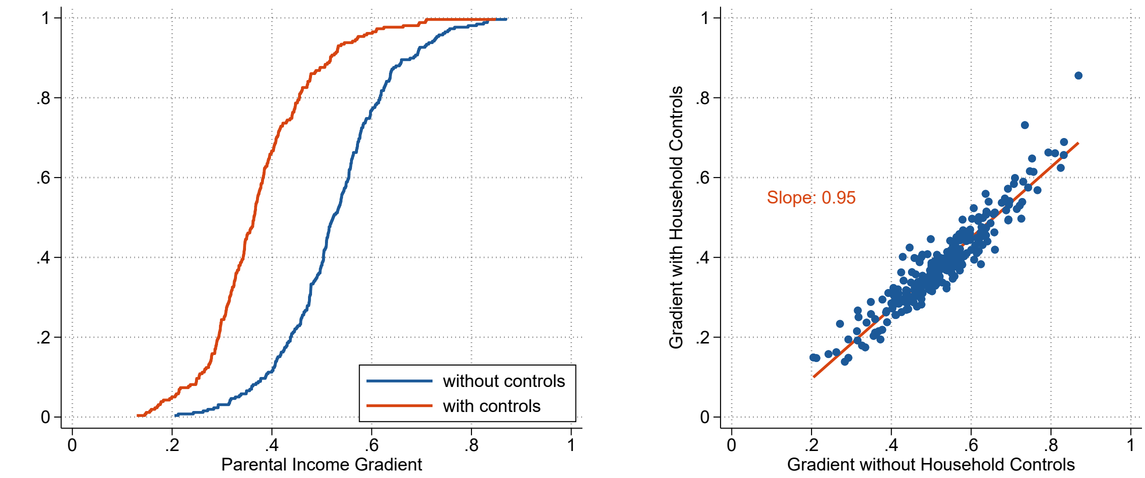
Our data, in contrast, contain a rich set of covariates and we can directly test to what extent differences in household characteristics can account for the geographical variation in mobility. To proceed, we recompute the parental income gradient for every LLM l but control for a comprehensive set of household characteristics X_i, including age and gender of the child, age and marital status of the parents, the number of siblings, a dummy for single parents and the highest parental education level in four categories. This results in the following model:

$$Y_{i,l} = \alpha_l + \beta_l R_l + \gamma X_i + \varepsilon_{i,l}. \tag{6}$$

We constrain the effects of additional household characteristics to be equal in all LLMs, as γ does not have a local-labor market specific subscript. By comparing the conditional parental income gradient ($\beta_l \times 100$) from this equation to the unconditional gradient estimated so far, we can assess the influence of household characteristics on geographical mobility differences. Figure 8, Panel (a) plots the Cumulative Distribution Function (CDF) of both gradients across all LLMs. The CDF of the unconditional gradient first order stochastically dominates the CDF of the conditional gradient, showing that in all LLMs the parental income gradient is reduced when controlling for household characteristics. This is intuitive, as especially controlling for parental education reduces the association between our outcome of interest and the parental income rank (compare Figure 2).

At the same time, the variance of the distribution (second moment) is not reduced when moving from the unconditional to the conditional gradient and, critically, the ranking between LLMs is largely unaffected: Figure 8, Panel (b) shows a scatter plot of the unconditional and the conditional gradients. The estimated linear fit of 0.95 demonstrates that controlling for household characteristics affects the estimates in most LLMs equally. The correlation coefficient is 0.91 and the Spearman rank correlation

Figure 8. Adjustment for Household Characteristics



(a) Cumulative Distribution Function

(b) Scatter Plot

Notes: Panel (a) plots the Cumulative Distribution Function (CDF) for the parental income gradient estimated with and without household controls across LLMs in Germany. Panel (b) shows a scatter plot of LLM-specific parental income gradients estimated when not controlling for additional household characteristics (x-axis) and LLM-specific Parental Income Gradients estimated when controlling for additional household characteristics (y-axis).

0.89. We conclude that differences in household characteristics seem to explain only a small fraction in mobility differences between LLMs.

7.2 Place-Based Predictors of Mobility

As household characteristics cannot account for the large spatial variation in mobility, it must be that place-based characteristics play an important role in shaping these differences. While our setting does not allow to identify their causal determinants, we can examine which regional characteristics are most predictive for spatial differences in mobility. This can provide a first indication of what factors may matter for mobility and guide future research. For this aim, we collect a comprehensive database of 71 regional indicators with information on labor market participation, economic conditions, infrastructure, demographics, housing and living conditions, the education structure and social characteristics. Appendix F provides more details on these regional indicators.

To study the association between local characteristics and intergenerational mobility, prior literature has typically relied on correlation coefficients or estimated multiple linear models via OLS (Chetty et al., 2014; Corak, 2020). Both approaches have disadvantages. As socio-economic characteristics are highly correlated at the regional level, correlation coefficients are often spurious. While this remedy is overcome in a multiple linear OLS regression, these models are prone to overfitting in high-dimensional data sets (Babyak, 2004). This may hamper the generalizability of the estimates. One way to address this it to reduce dimensionality of the covariates via variable selection. Belloni and Chernozhukov (2013) suggest to preselect covariates via Lasso before estimating a multiple linear model.[19] This approach is for example applied by Finkelstein et al. (2016) to explain geographical variation in health care utilization in the US.

We take a similar two-step approach, but preselect variables using a Random Forest variable importance measure instead of preselecting variables with a Lasso regression. This is because we find that a linear Lasso model only poorly fits our data due to non-linear interactions between the variables.[20] The Random Forest algorithm in contrast is fully non-parametric and can capture higher-order interactions in the data. After fitting the Random Forest, we can rank covariates according to their predictive power and thus obtain a measure of variable importance.[21]

Variable Selection. In a first step, we therefore use a Random Forest to predict the parental income gradient for every LLM based on all 71 regional indicators. We then compute a measure of variable importance and rank predictors accordingly. The 15 most informative predictors are displayed in Table 7.[22] Overall, mainly social characteristics, the local organization of the education system and labor market conditions seem to matter for mobility. The column to the right displays the sign of the bivariate correlation between each variable and the parental income gradient. A positive

[19] An alternative approach to deal with model uncertainty is model averaging. See Kourtellos et al. Kourtellos, Marr, and Tan (2016) for an application in the context of social mobility.

[20] To compare the out-of-sample performance of this algorithm against an implementation of a Lasso and an Elastic Net regression with α =0.5, we split our data in a training and test data set (75-25 split). The Random Forest algorithm predicts 38% of the variation in the test sample (R^2 = 0.38), whereas the predictive power of Lasso (R^2 = 0.11) and Elastic Net (R^2 = 0.17) is much lower. The results for Lasso and Elastic Net are are based on λ chosen by 5-fold cross-validation. For the Random Forest, we fit 1000 trees and randomly select 72/3 = 24 variables for each split.

[21] There are several ways to compute a Random Forest based measure of variable importance. We choose the implementation proposed by Strobl et al. (2008), which computes a conditional permutation importance measure that accounts for the dependence structure between the predictors. The Random Forest variable importance measure used here is hence conceptually similar to a variable importance ranking provided by Lasso - with the addition that it is fully flexible to account for non-linear interactions between covariates.

[22] The exact ranking of predictors especially after rank 5 varies for different implementations of the Random Forest algorithm. We are therefore cautious not to over-interpret the ranking between single predictors. As described above, this ranking only serves to reduce dimensionality of our covariates.

Table 7. The 15 Most Predictive Regional Indicators

Rank	Variable	Importance Measure	ρ
1	School Dropout Rate	1.043e-07	+
2	Share Married	5.349e-08	−
3	Students	4.551e-08	−
4	Teenage Pregnancies	2.615e-08	+
5	Broadband Availability	1.715e-08	+
6	Distance to Next College	1.465e-08	−
7	Median Income Vocational Qualification	1.225e-08	−
8	Unemployment Rate	1.207e-08	+
9	Gender Wage Gap	1.204e-08	+
10	Share without Vocational Qualification	9.343e-09	−
11	Gini Parental Income	6.757e-09	−
12	Share on Vocational A-Level Track	6.356e-09	−
13	Voter Turnout	5.134e-09	−
14	Child Poverty	4.061e-09	+
15	Share Children 0-2 in Childcare	4.052e-09	+

Notes: This table lists the 15 most predictive indicators for explaining variation in the parental income gradient between LLMs in Germany. See text for details on the implementation via a Random Forest variable importance measure. The last column shows the sign of the Pearson correlation coefficient between each variable and the parental income gradient. A positive coefficient therefore implies that the variable is predictive of a low degree of social mobility.

sign implies that the indicator is predictive for low mobility (a high gradient). For example, LLMs with a high prevalence of school dropouts are associated with low relative mobility. The same applies to the share of teenage pregnancies, the prevalence of child poverty and the share of individuals which are dependent on social assistance. All these indicators point to comparatively disadvantaged social contexts in these labor markets, consistent with social capital based explanations of regional disparities in mobility. A high share of married individuals in contrast signals high mobility. Other variables like the access to broadband Internet or the distance to the next elementary school are less straightforward to interpret, but apparently predictive for mobility.

Regression Estimates. In a second step, we regress the gradient on all 15 indicators. All right-hand side variables are standardized so that the coefficients report the association between a one standard deviation change in the covariate and the absolute change in the gradient. The results are reported in Table 8. The signs of the coefficients generally match those from the bivariate correlations in Table 7. For example, a one standard deviation increase in the school dropout rate is associated with a

Table 8. Social Mobility and Regional Characteristics

	(1)	(2)	(3)	(4)	(5)
School Dropout Rate	0.0450	0.0442	0.0419	0.0550	0.0543
	(0.0112)	(0.0114)	(0.0095)	(0.0158)	(0.0158)
Share Married	-0.0195	-0.0267	-0.0188	-0.0165	-0.0206
	(0.0084)	(0.0085)	(0.0063)	(0.0102)	(0.0103)
Students	-0.0228	-0.0248	-0.0117	-0.0271	-0.0301
	(0.0128)	(0.0126)	(0.0094)	(0.0160)	(0.0159)
Teenage Pregnancies	0.0305	0.0266	0.0292	0.0249	0.0194
	(0.0234)	(0.0233)	(0.0172)	(0.0261)	(0.0271)
Broadband Availability	0.0218	0.0236	0.0201	0.0203	0.0230
	(0.0108)	(0.0107)	(0.0090)	(0.0111)	(0.0112)
Distance to Next College	-0.0043	-0.0063	-0.0071	-0.0048	-0.0079
	(0.0074)	(0.0078)	(0.0073)	(0.0075)	(0.0080)
Median Income Vocational	-0.0168	-0.0144	0.0046	-0.0171	-0.0138
Qualification	(0.0114)	(0.0115)	(0.0114)	(0.0142)	(0.0151)
Unemployment Rate	0.0168	0.0144	0.0064	0.0271	0.0201
	(0.0327)	(0.0324)	(0.0239)	(0.0416)	(0.0417)
Gender Wage Gap	-0.0045	-0.0065	0.0023	0.0061	0.0005
	(0.0143)	(0.0144)	(0.0132)	(0.0177)	(0.0182)
Share without Vocational	0.0062	0.0067	-0.0096	0.0176	0.0158
Qualification	(0.0162)	(0.0162)	(0.0117)	(0.0220)	(0.0222)
Gini Parental Income	-0.0222	-0.0187	-0.0297	-0.0051	0.0013
	(0.0143)	(0.0144)	(0.0105)	(0.0196)	(0.0205)
Share on Vocational	-0.0159	-0.0160	-0.0131	-0.0201	-0.0207
A-Level Track	(0.0093)	(0.0091)	(0.0082)	(0.0102)	(0.0101)
Voter Turnout	0.0199	0.0228	0.0092	0.0254	0.0314
	(0.0122)	(0.0123)	(0.0095)	(0.0168)	(0.0172)
Child Poverty	-0.0415	-0.0260	-0.0350	-0.0496	-0.0223
	(0.0319)	(0.0332)	(0.0239)	(0.0385)	(0.0444)
Share Children 0-2	-0.0440	-0.0430	-0.0494	-0.0351	-0.0396
in Childcare	(0.0183)	(0.0181)	(0.0184)	(0.0261)	(0.0261)
Additional Controls	-	✓	✓	-	✓
State Fixed Effects	-	-	-	✓	✓
Weighted	-	-	✓	-	-
N	258	258	258	252	252
R^2	0.259	0.279	0.242	0.293	0.306

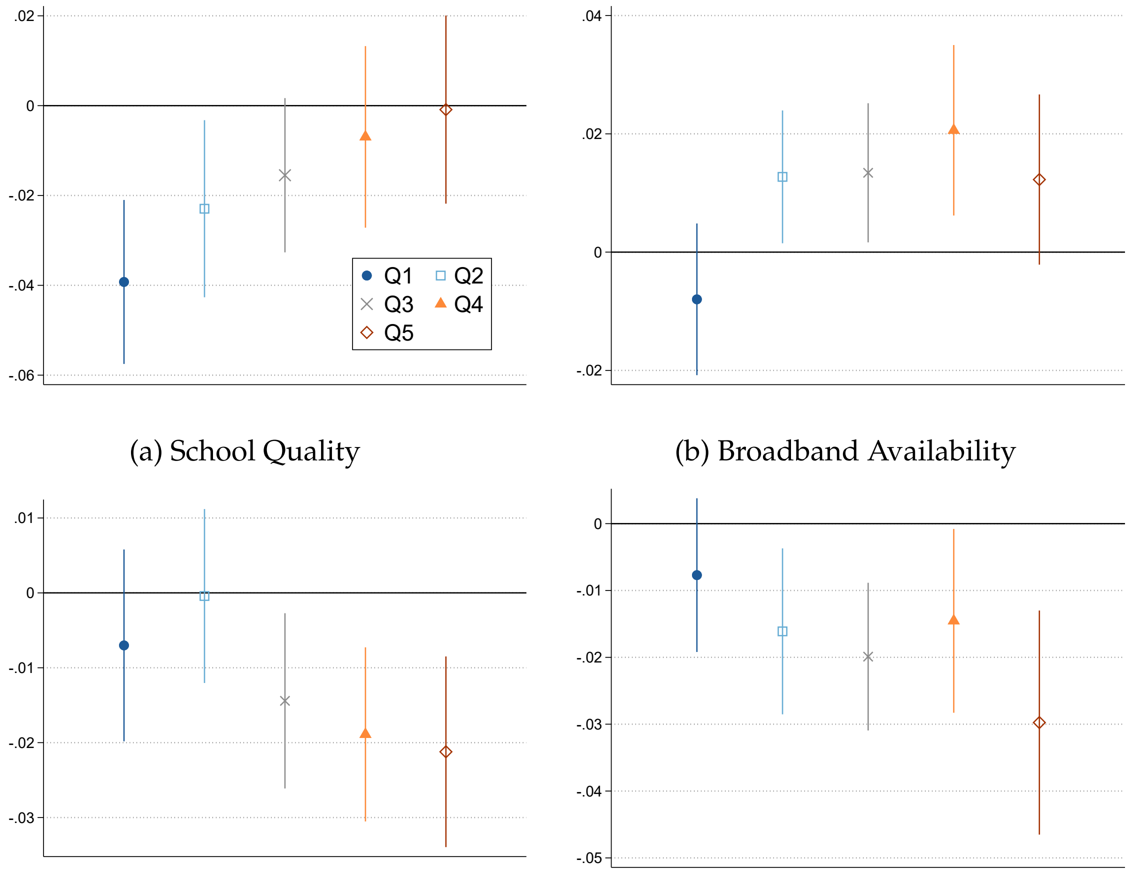
Notes: Each column of this table reports coefficients from an linear regression with robust standards errors reported in parentheses. The dependent variable in all columns is the parental income gradient. The independent variables (as selected by the Random Forest, compare Table 7) are standardized to have mean 0 and standard deviation 1. Columns (3) and (4) contain state fixed effects, for which we have to drop five LLMs crossing state borders and the LLM of Berlin. In columns (2) and (4), we additionally control for population, population density and the region type (rural, urban or mixed) to test if coefficients of the regional indicators are affected by structural differences in mobility between more rural or urban LLMs. In column (3) we weight the regression with the number of observations per LLM.

4.5 percentage point higher parental income gradient. This association becomes even stronger when adding state fixed effects. A high gradient also aligns with a high number of teenage pregnancies, enhanced child poverty, a higher unemployment rate, a high voter turnout and a high share of households with broadband Internet access. A negative association with the parental income gradient arises for the share of married individuals, the distance to the next college, the median income for individuals with a recognized vocational qualification, the share of children on a vocational A-Level track and for the share of children aged 0-2 in childcare. However, due to the limited sample size, we lack the power to precisely estimate most coefficients. Exceptions are the school dropout rate, the broadband availability, the share of married individuals and the share of children on a vocational A-Level track.

Graphical Evidence. To better understand the relationship between relative mobility and these statistically significant indicators, we separately regress the A-Level share in each quintile of the parental income distribution on each indicator and plot the estimates in Figure 9. These plots reveal whether, for example, a positive relationship between the parental income gradient and an indicator is driven by a lower A-Level share of children from low-income households or by a higher A-Level share of children from high-income households.

We start with the school dropout rate. In the US context, Chetty et al. (2014) interpret the school dropout rate – adjusted by parental income – as an indicator of school quality and find a strong negative correlation with relative mobility. In close analogy, we regress the dropout rate on mean parental income, the Gini coefficient of parental income, the share of parents holding an A-Level degree and the unemployment rate and take the residuals to obtain a measure of school quality which is adjusted for parental income and labor market conditions. This indicator is still highly correlated with mobility. As depicted in Figure 9, Panel (a), low school quality (a high value of the indicator) is associated with a lower probability to obtain an A-Level degree for children from low income households but does not seem to affect children in the top two quintiles of the parental income distribution. While this would be consistent with the idea that school quality is crucial for improving opportunities for children from

Figure 9. Predicting the A-Level Share by Parental Income Quintile



(a) School Quality

(b) Broadband Availability

(c) Share Married

(d) Share on Vocational A-Level Track

Notes: Each panel of this figure reports coefficients from five separate linear OLS regressions with robust standards errors and 95% confidence bands. The dependent variable is the share of children which obtained an A-Level in the respective quintile of the parental income distribution. The independent variable is the adjusted school dropout rate (school quality index) in Panel (a), the share of broadband connections per 100 inhabitants in Panel (b), the share of married individuals in Panel (c) and the share of students on a vocational (rather than general education) A-Level track (Panel d). In addition, all regression include a set of state fixed effects and control for population, population density and the region type (rural, urban or mixed). We exclude 6 LLMs with insufficient observations for estimating Q5 from the sample. Due to the inclusion of state fixed effects, we have to further drop five LLMs crossing state borders and the LLM of Berlin from the sample, leaving us with 246 observations. All regressors are standardized to have mean 0 and standard deviation 1.

low socio-economic background, further information is needed to test this hypothesis in detail.[23]

[23]Most importantly, it remains open if the adjusted school drop out rate is indeed an appropriate proxy for school quality. In the US, Rothstein (2019) studies how closely the transmission of parental income to educational attainment and achievement (test scores) are correlated with income mobility at the commuting zone level. He finds income-income transmission to be closely connected to income-educational attainment transmission but not to income-educational achievement transmission. Rothstein (2019) therefore finds little evidence that differences in the quality of secondary schooling are a key mechanism driving variation in intergenerational mobility. However, the distinct features of the German secondary schooling system could lead to very different patterns in our data. Unfortunately,

Panel (b) sheds light on the negative connection between broadband availability and mobility. While broadband access is associated with a higher A-Level share on average, this is not true for children in Q1, for whom the relationship becomes negative. We can only speculate about the reasons. Broadband access is highly correlated with factors pointing at dynamic and prosperous labor markets which exhibit above average inequality. For that reason, broadband availability may proxy urban areas in which most children profit from a dynamic and rewarding economic environment, but the kids at the bottom fail to take part in this development. However, broadband availability could also causally influence social mobility. For the US, Dettling et al. (2018) document that increased broadband availability fosters access to college and find the effect to be concentrated among students with parents from high socio-economic status. Similarly, Sanchis-Guarner et al. (2021) report a causal (positive) impact of broadband access on student test scores in England but find comparatively lower effects for students eligible for free school meals. Our results would be in line with these findings.

The opposite pattern emerges for the share of married individuals in Panel (c): this statistic is related to higher mobility but a lower A-Level share of children from high-income families. Finally, Panel (d) reports the association between the Q-measures and the share of children on a vocational, rather than general interest, A-Level track. There is reason to believe that the availability of such vocational tracks may dampen the influence of parental background for the opportunities of children. Children in these tracks have typically obtained a degree from the medium track (Realschule) and now attend a specialized vocational school to obtain an A-Level degree on top. In that setting, vocational schools may especially foster the opportunities of children from low-income households initially "misallocated" to the medium instead of the high track. Dustmann et al. (2017) show that vocational schools have the potential to fully offset adverse affects of early age tracking on long-term labor market outcomes, but cannot observe parental background.

Our evidence shows that, relative to children from the top quintile, children from the bottom quintile are more likely to obtain an A-Level in LLMs with a high prevalence of such schools. In addition, we find that at the national level the parental income rank is more predictive for the probability to attend the general high track (Gymnasium) at the age of 13-14 than to obtain an A-Level degree later on (gradient of 0.55

there exist no comparable data on student test scores in Germany, preventing us from investigating this issue further.

versus 0.52), again suggesting that vocational schools may mediate the influence of parental background.

8 Conclusion

In this paper, we use census data to characterize social mobility in Germany at a higher level of detail than previously possible. We measure mobility by the association between the educational attainment of a child and the parents' percentile rank in the income distribution. At the national level, a 10 percentile increase in the parental income rank is associated with a 5.2 percentage point increase in the probability to obtain an A-Level degree. This relationship remained stable for recent cohorts, despite a massive roll-out of higher secondary education in the period 1997-2018. An expansion in access to higher education alone may therefore not be sufficient to reduce the opportunity gap between children from high and low income households.

At the same time, mobility varies substantially across areas, mainly within states. We show that place-based characteristics, rather than sorting of households into different regions, explain most of these differences. We find that social characteristics, the local organization of the education system and labor market conditions best predict mobility at the regional level. More research is needed to understand whether these correlations reflect structural relationships. The mobility statistics constructed here could serve as a starting point for this kind of analysis, as they provide variation in both time and space. Arguably exogenous events like state-level reforms of the education system or local labor market specific shocks could be exploited to shed light on the determinants of mobility.

The approach described in this paper provides a timely and feasible way to monitor the development of social mobility in Germany for recent cohorts. This measurement framework may also prove useful in other countries where the highest secondary school degree is crucial for future career options. Education systems with secondary school degrees of comparable importance to the Abitur in Germany include Italy (Maturità), Austria (Matura) and the UK (A-Level). If similar census data were available for these countries, it would be interesting to see how estimates compare to the findings provided in this paper.

References

ACCIARI, P., A. POLO, AND G. L. VIOLANTE (2019): “"And Yet, It Moves": Intergenerational Mobility in Italy,” *NBER Working Paper No. 25732*.

ALESINA, A., S. HOHMANN, S. MICHALOPOULOS, AND E. PAPAIOANNOU (2021): “Intergenerational Mobility in Africa,” *Econometrica*, 89(1), 1–35.

ASHER, S., P. NOVOSAD, AND C. RAFKIN (2020): “Intergenerational Mobility in India: New Methods and Estimates Across Time, Space, and Communities,” *http://paulnovosad.com/pdf/anr-india-mobility.pdf*.

BABYAK, M. A. (2004): “What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models,” *Psychosomatic Medicine*, 66(3), 411–421.

BELLONI, A., AND V. CHERNOZHUKOV (2013): “Least Squares after Model Selection in High-Dimensional Sparse Models,” *Bernoulli*, 19(2), 521–547.

BIEWEN, M., AND M. TAPALAGA (2017): “Life-Cycle Educational Choices in a System with Early Tracking and ‘Second Chance’ Options,” *Economics of Education Review*, 56, 80–94.

BRATBERG, E., J. DAVIS, B. MAZUMDER, M. NYBOM, D. SCHNITZLEIN, AND K. VAAGE (2017): “A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US,” *Scandivian Journal of Economics*, 119(1), 72–101.

CHETTY, R., AND N. HENDREN (2018a): “The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects,” *Quarterly Journal of Economics*, 133(3), 1107–1162.

——— (2018b): “The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates,” *Quarterly Journal of Economics*, 133(3), 1163–1228.

CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (2014): “Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States,” *Quarterly Journal of Economics*, 129(4), 1553–1623.

CHUARD, P., AND V. GRASSI (2020): "Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education," *University of St. Gallen, Department of Economics Discussion Paper No. 2020-11*.

CORAK, M. (2020): "The Canadian Geography of Intergenerational Income Mobility," *Economic Journal*, 130(631), 2134–2174.

DAHL, M., AND T. DELEIRE (2008): *The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data*. University of Wisconsin-Madison, Institute for Research on Poverty.

DAHRENDORF, R. (1966): *Bildung ist Bürgerrecht. Plädoyer für eine aktive Bildungspolitik.* Nannen-Verlag, Hamburg.

DETTLING, L. J., S. GOODMAN, AND J. SMITH (2018): "Every Little Bit Counts: The Impact of High-Speed Internet on the Transition to College," *Review of Economics and Statistics*, 100(2), 260–273.

DEUTSCHER, N., AND B. MAZUMDER (2020): "Intergenerational Mobility across Australia and the Stability of Regional Estimates," *Labour Economics*, 66, 101861.

DUSTMANN, C., P. PUHANI, AND U. SCHÖNBERG (2017): "The Long-Term Effects of Early Track Choice," *Economic Journal*, 127, 1348–1380.

EBERHARTER, V. V. (2013): "The Intergenerational Dynamics of Social Inequality – Empirical Evidence from Europe and the United States," *SOEPpapers 588-2013*.

EISENHAUER, P., AND F. PFEIFFER (2008): "Assessing Intergenerational Earnings Persistence among German Workers," *Journal of Labour Market Research*, 2/3, 119–137.

FINKELSTEIN, A., M. GENTZKOW, AND H. WILLIAMS (2016): "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *Quarterly Journal of Economics*, 131(4), 1681–1726.

GALLAGHER, R., R. KAESTNER, AND J. PERSKY (2018): "The Geography of Family Differences and Intergenerational Mobility," *Journal of Economic Geography*, 19(3), 589–618.

GÄRTNER, K. (2002): "Differentielle Sterblichkeit. Ergebnisse des Lebenserwartungssurvey des BiB.," *Zeitschrift für Bevölkerungswissenschaft*, 27(2), 185 – 211.

HADJAR, A., AND R. BECKER (2006): "Bildungsexpansion – Erwartete und Unerwartete Folgen," in *Die Bildungsexpansion*, pp. 11–24. VS Verlag für Sozialwissenschaften.

HAIDER, S., AND G. SOLON (2006): "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, 96(4), 1308–1320.

HAUSNER, K. H., D. SÖHNLEIN, B. WEBER, AND E. WEBER (2015): "Qualifikation und Arbeitsmarkt: Bessere Chancen mit mehr Bildung," *IAB Kurzbericht 11/2015*.

HILGER, N. G. (2015): "The Great Escape: Intergenerational Mobility in the United States Since 1940," *NBER Working Paper No. 21217*.

KLEIN, M., K. BARG, AND M. KÜHHIRT (2019): "Inequality of Educational Opportunity in East and West Germany: Convergence or Continued Differences?," *Sociological Science*, 6, 1–26.

KLINE, P., AND E. MORETTI (2014): "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, 6(1), 629–662.

KOURTELLOS, A., C. MARR, AND C. M. TAN (2016): "Robust Determinants of Intergenerational Mobility in the Land of Opportunity," *European Economic Review*, 81, 132–147.

LEE, C.-I., AND G. SOLON (2009): "Trends in Intergenerational Income Mobility," *Review of Economics and Statistics*, 91(4), 766–772.

LOCHNER, L. (2011): "Nonproduction Benefits of Education: Crime, Health, and Good Citizenship," *Handbook of the Economics of Education*, 4, 183–282.

MAZUMDER, B. (2005): "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data," *Review of Economics and Statistics*, 87(2), 235–255.

——— (2016): "Estimating the Intergenerational Elasticity and Rank Association in the U.S.: Overcoming the Current Limitations of Tax Data," *Research in Labor Economics*, 43, 83–129.

——— (2018): "Intergenerational Mobility in the United States: What we Have Learned from the PSID," *Annals of the American Academy of Political and Social Science*, 680(1), 213–234.

MUÑOZ, E. (2021): "The Geography of Intergenerational Mobility in Latin America and the Caribbean," *STONE CENTER ON SOCIO-ECONOMIC INEQUALITY WORKING PAPER SERIES No. 29*.

NEUMARK, D., AND H. SIMPSON (2015): "Place-Based Policies," in *Handbook of Regional and Urban Economics*, vol. 5, pp. 1197–1287. Elsevier.

NYBOM, M., AND J. STUHLER (2016): "Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation," *Journal of Human Resources*, 51(1), 239–268.

——— (2017): "Biases in Standard Measures of Intergenerational Income Dependence," *Journal of Human Resources*, 52(3), 800–825.

OREOPOULOS, P., AND K. G. SALVANES (2011): "Priceless: The Nonpecuniary Benefits of Schooling," *Journal of Economic Perspectives*, 25(1), 159–84.

PICHT, G. (1964): *Die deutsche Bildungskatastrophe: Analyse und Dokumentation*. Walter Verlag.

RIPHAHN, R., AND G. HEINECK (2009): "Intergenerational Transmission of Educational Attainment in Germany – The Last Five Decades," *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, 229(1), 36–60.

RIPHAHN, R., AND P. TRÜBSWETTER (2013): "The Intergenerational Transmission of Educational Attainment in East and West Germany," *Applied Economics*, 45(22), 3183–3196.

ROTHBAUM, J. (2016): "Sorting and Geographic Variation in Intergenerational Mobility," *Unpublished Manuscript*.

ROTHSTEIN, J. (2019): "Inequality of Educational Opportunity? Schools as Mediators of the Intergenerational Transmission of Income," *Journal of Labor Economics*, 37(1), 85–123.

SANCHIS-GUARNER, R., J. MONTALBÁN, AND F. WEINHARDT (2021): "Home Broadband and Human Capital Formation," *CESifo Working Paper No. 8846*.

SCHMILLEN, A., AND H. STÜBER (2014): "Lebensverdienste nach Qualifikation - Bildung Lohnt sich ein Leben Lang," *IAB Kurzbericht 1/2014*.

SCHNITZLEIN, D. (2016): "A New Look at Intergenerational Mobility in Germany Compared to the US," *Review of Income and Wealth*, 62(4), 650–667.

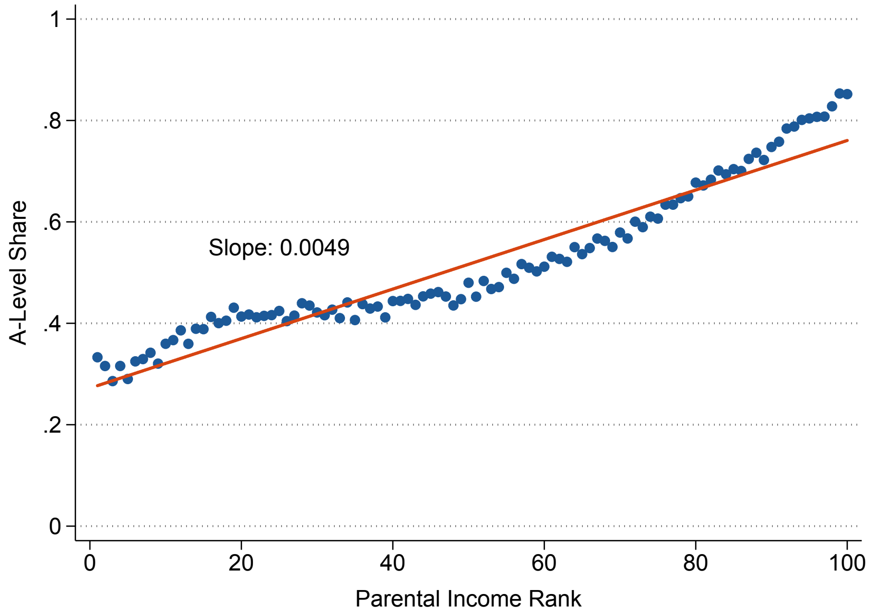
SOLON, G. (1992): "Intergenerational Income Mobility in the United States," *American Economic Review*, 82(3), 393–408.

STROBL, C., A.-L. BOULESTEIX, T. KNEIB, T. AUGUSTIN, AND A. ZEILEIS (2008): "Conditional Variable Importance for Random Forests," *BMC Bioinformatics*, 9(1), 307.

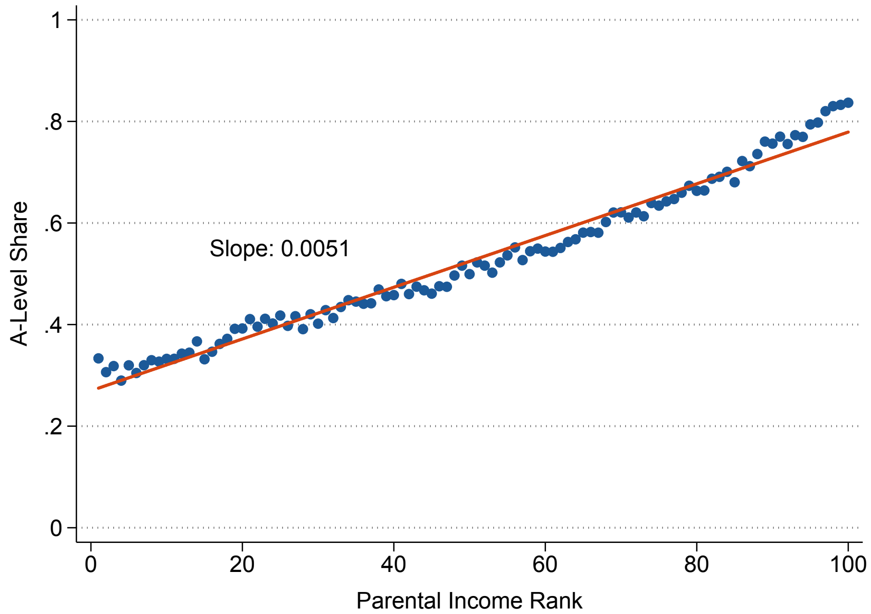
ZIMMERMAN, D. J. (1992): "Regression Toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–429.

A Alternative Equivalence Scales

Figure A.1. National Estimates with Different Equivalization Schemes



(a) No Adjustment

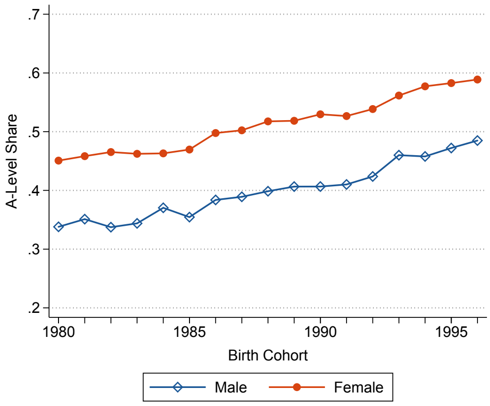


(b) Per Capita Adjustment

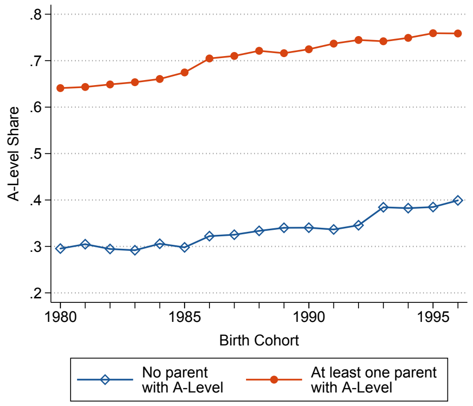
Notes: This figure shows non-parametric binned scatter plots of the fraction of children aged 17-21 that are either enrolled in the last two/three years of the A-Level track or already completed the A-Level degree by percentile rank of their parents in the national income distribution in the period 2011-2018. In Panel (a), parental income is not adjusted for household size, whereas in Panel (b) we adjust income by dividing through the household size. The parental percentile rank is computed within each year and the OLS slopes reported in the figure are estimated using the underlying micro data.

B Trends for Subgroups

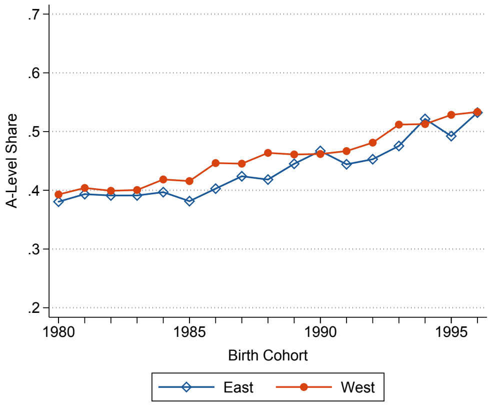
Figure B.1. Time Trend A-Level Share for Subgroups



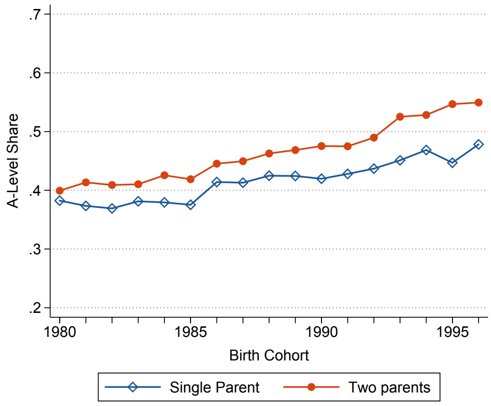
(a) Gender



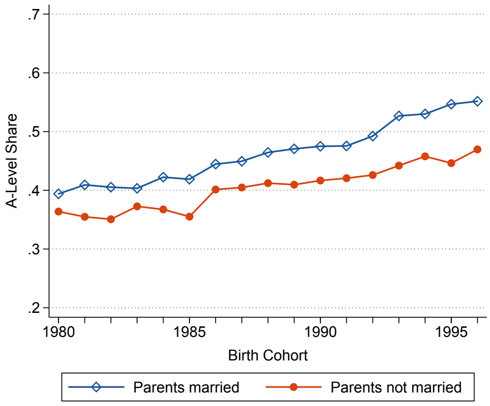
(b) Parental Education



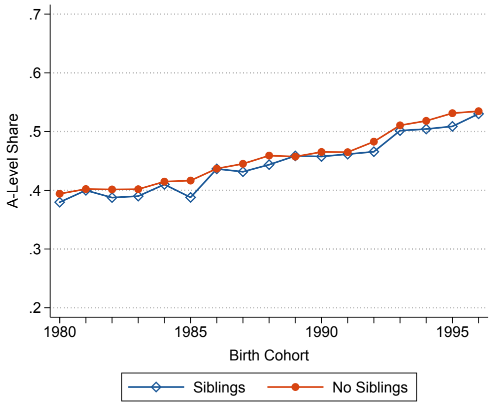
(c) Region



(d) Parenting Status



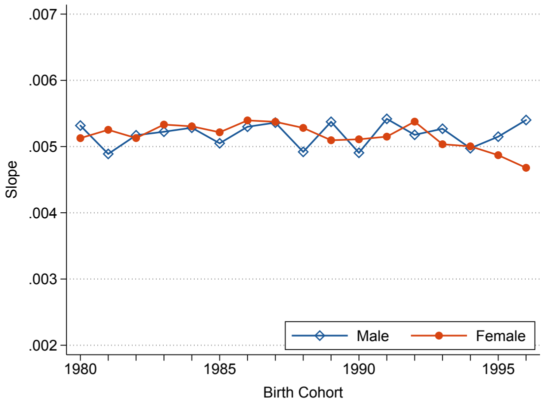
(e) Parental Marital Status



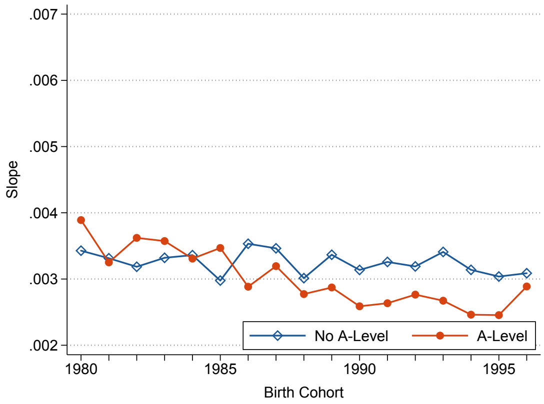
(f) Siblings

Notes: This figure shows the development of the A-Level share for different population subgroups for birth cohorts 1980-1996 in the MZ. The A-Level share is given as the fraction of children aged 17-21 that are either enrolled in the last two/three years of the A-Level track or already completed an A-Level degree.

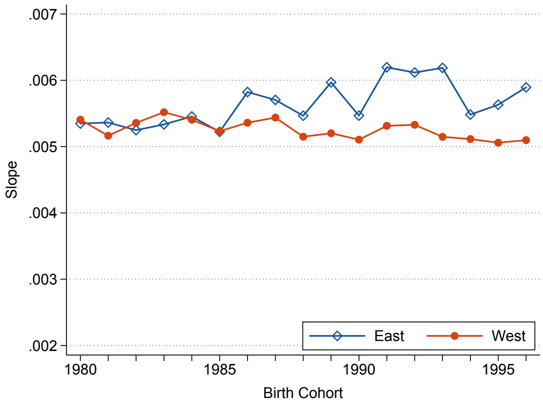
Figure B.2. Time Trend Parental Income Gradient for Subgroups



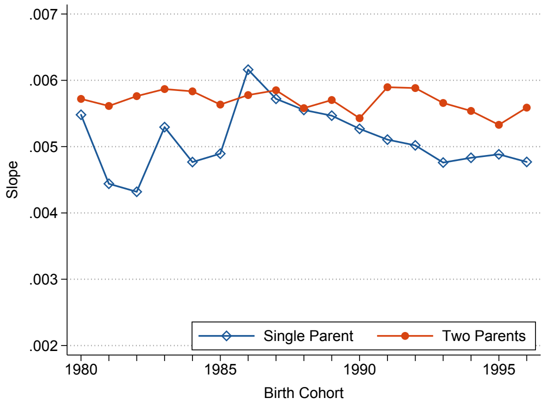
(a) Gender



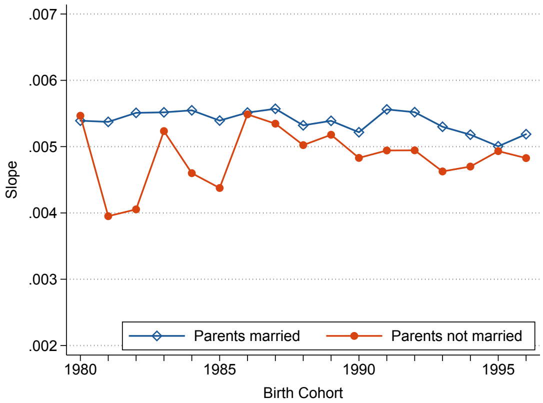
(b) Parental Education



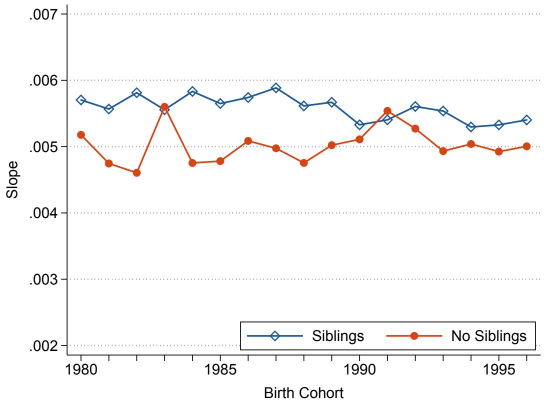
(c) Region



(d) Parenting Status



(e) Parental Marital Status

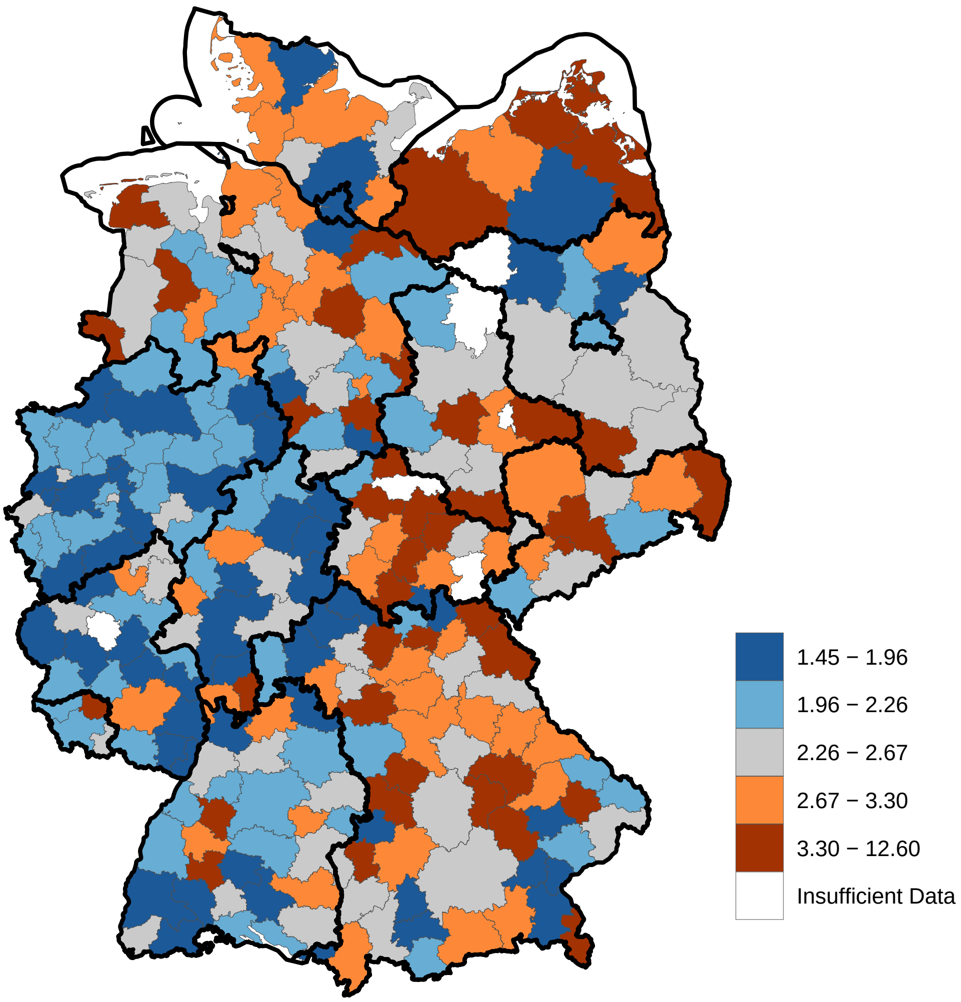


(f) Siblings

Notes: This figure displays the development of relative mobility for different population subgroups for birth cohorts 1980-1996 in the MZ. We display the slope coefficients γ_t of the following regression: $Y_{i,t} = \alpha + \beta_t C_t + \gamma_t C_t \times R_i + \varepsilon_{i,t}$, where C_t denotes a cohort and $C_t \times R_i$ the interaction between cohort and parental income rank. The slope coefficient γ_t multiplied by 100 uncovers the parental income gradient, separately for each cohort.

C Heatmap of the Q5/Q1 Ratio

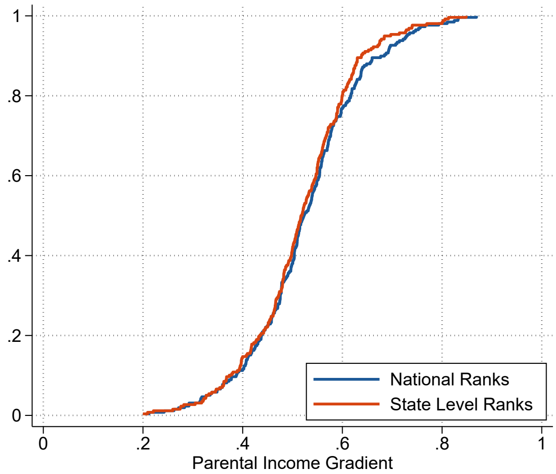
Figure C.1. Q5/Q1 Ratio by Local Labor Market



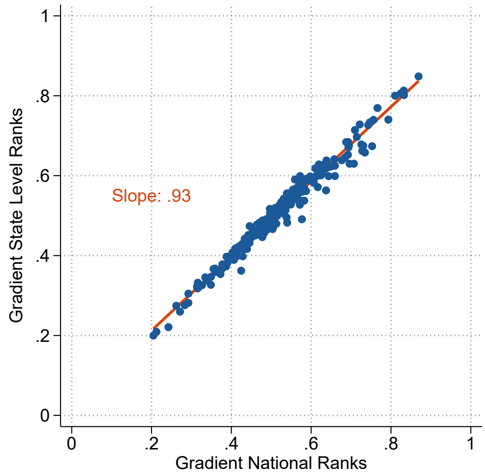
Notes: This figure presents a heat map of the Q5/Q1 ratio by LLM. Children are assigned to LLMs according to their current place of residence. The estimates are based on children aged 17-21 in the years 2011-2018 for which we have non-missing information on educational attainment and parental income. The Q5/Q1 ratio is computed by dividing the share of children with an A-Level degree in the top 20% through the share of children with an A-Level degree in the bottom 20% of the parental income distribution. 6 LLMs with less than three children in the top 20% of the parental income distribution without an A-Level degree are excluded from the analysis.

D Robustness Geography of Mobility

Figure D.1. State Specific Parental Income Ranks



(a) Cumulative Distribution Function



(b) Scatter Plot

Notes: This figure compares the distribution of the parental income gradient across LLMs in Germany when the parental income rank is computed either in the national or the state-specific income distribution. Panel (a) displays the Cumulative Distribution Function (CDF) of both gradients and shows that both mean and variance vary only marginally. Panel (b) plots the parental income gradient with ranks computed in the national income distribution (y-axis) against the parental income gradient with ranks computed in the state-specific income distribution (x-axis). The OLS slope of 0.93 implies a very high correlation between both estimates, showing that the distribution of the parental income gradient across LLMs is largely insensitive to the chosen reference income distribution.

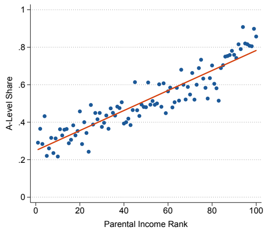
E Mobility at the State Level

Table E.1. Mobility at the State Level

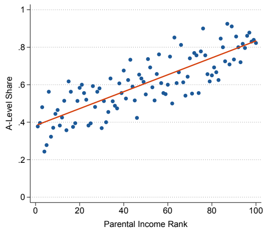
Rank	Name	Gradient	Q5/Q1	Q1	A-Level Share
1	Hamburg	0.45	1.85	0.43	0.60
2	Rheinland-Pfalz	0.50	2.11	0.36	0.53
3	Nordrhein-Westfalen	0.51	2.02	0.41	0.59
4	Hessen	0.52	2.07	0.39	0.59
5	Baden-Württemberg	0.52	2.25	0.34	0.53
6	Saarland	0.53	2.27	0.33	0.54
7	Schleswig-Holstein	0.53	2.35	0.32	0.52
8	Niedersachsen	0.54	2.54	0.29	0.48
9	Bayern	0.55	2.72	0.25	0.42
10	Berlin	0.56	2.19	0.39	0.59
11	Brandenburg	0.57	2.35	0.36	0.60
12	Sachsen-Anhalt	0.57	2.82	0.25	0.43
13	Sachsen	0.61	2.86	0.27	0.48
14	Mecklenburg-Vorpommern	0.63	3.01	0.25	0.45
15	Bremen	0.64	2.65	0.32	0.55
16	Thüringen	0.64	3.07	0.25	0.46

Notes: This table reports estimates of absolute and relative mobility for all 16 German states. States are sorted in descending order by the relative mobility rank, as measured by the parental income gradient. A higher rank thus indicates higher relative mobility.

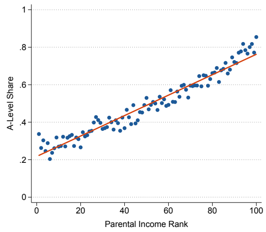
Figure E.1. State Level Gradients



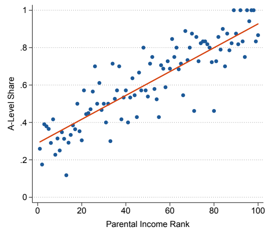
(a) Schleswig-Holstein



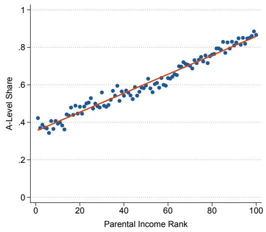
(b) Hamburg



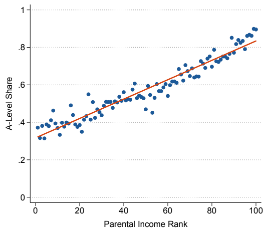
(c) Niedersachsen



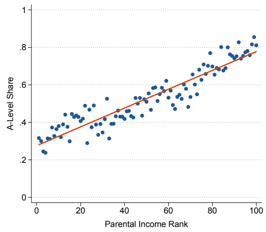
(d) Bremen



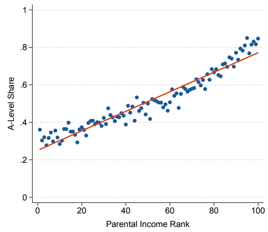
(e) Nordrhein-Westfalen



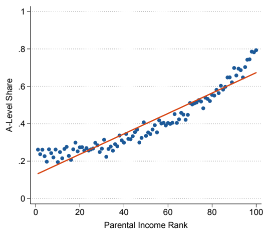
(f) Hessen



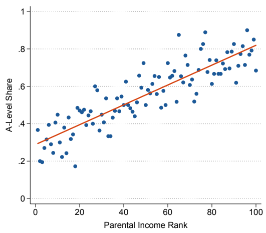
(g) Rheinland-Pfalz



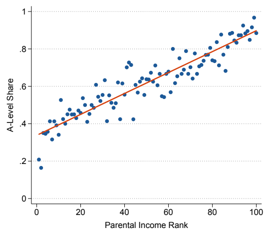
(h) Baden-Württemberg



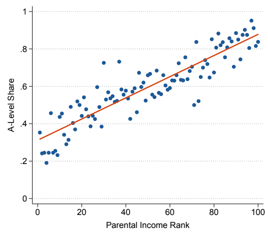
(i) Bayern



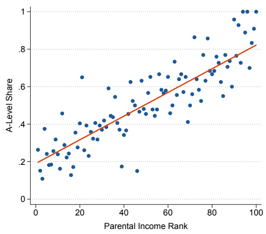
(j) Saarland



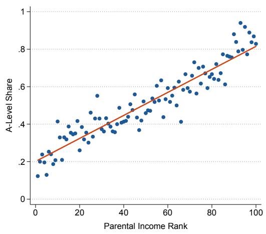
(k) Berlin



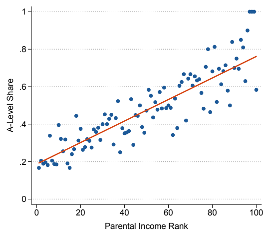
(l) Brandenburg



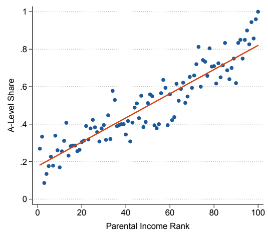
(m) Mecklenburg-Vorpommern



(n) Sachsen



(o) Sachsen-Anhalt



(p) Thüringen

Notes: This figure shows for each German state a binned scatter plot of the fraction of children aged 17-21 that are either enrolled in the last two/three years of the A-Level track or already completed the A-Level degree by percentile rank of their parents in the national income distribution in the period 2011-2018. The parental income rank is computed within each year.

F Details on Regional Indicators

Table F.1. List of Regional Indicators

Category	Variable	Source
Labor Market	Unemployment Rate	INKAR
	Share Long Term Unemployed	INKAR
	Share Female Employees	INKAR
	Share Part Time Employees	INKAR
	Share without Vocational qualification	INKAR
	Share Marginal Employment	INKAR
	Share Employed in Manufacturing Sector	INKAR
	Apprenticeship Positions	INKAR
	Apprentices	INKAR
	Vocational School Students	INKAR
	Employees with Academic Degree	INKAR
	Commuting Balance	INKAR
	Hours Worked	INKAR
	A-Level Wage Premium	MZ
Education	Students (before Tertiary Education)	INKAR
	Students (Tertiary Education)	INKAR
	Students (Universities of Applied Sciences)	INKAR
	School Dropout Rate	INKAR
	Highly Qualified Persons	INKAR
	Share Children 0-2 in Childcare	INKAR
	Share Children 3-5 in Childcare	INKAR
	Share of all Students Enrolled in Gymnasium	INKAR
	Share of all Secondary School Students Enrolled in Gymnasium	Destatis
	Distance to Next College	HRK
	Distance to Next Elementary School	INKAR
	Share on Vocational A-Level Track	MZ
	Share of Parents with an A-Level	MZ
Income	Median Household Income	INKAR
	Median Household Income with Vocational Qualification	INKAR
	Gender Wage Gap	INKAR
	Child Poverty	INKAR
	Mean Household Income	INKAR
	Gini Household Income	MZ
	Mean Parental Income	MZ
	Gini Parental Income	MZ
	Ratio p85/p50 (Household Income)	MZ
	Ratio p50/p15 (Household Income)	MZ

Economy	GDP per Capita	INKAR
	Municipal Tax Revenues per Capita	INKAR
	Municipal Debt per Capita	INKAR
	Business Creation	INKAR
Housing	Construction Land Prices	INKAR
	New Apartments	INKAR
	Building Permits	INKAR
	Living Area	INKAR
	Share of Apartment Buildings	INKAR
	Rent Prices	INKAR
Infrastructure	Physician Density	INKAR
	Broad Band Availability	INKAR
	Passenger Car Density	INKAR
	Hospital Beds	INKAR
Demographics	Average Age	INKAR
	Share Female	INKAR
	Share Foreigners	INKAR
	Share Asylum Seekers	INKAR
	Total Net Migration	INKAR
	Births Net of Deaths	INKAR
	Fertility Rate	INKAR
	Teenage Pregnancies	INKAR
	Life Expectancy	INKAR
	Child Mortality	INKAR
	Population Density	INKAR
	Share Single Parents	MZ
	Share Married	MZ
	Share Divorced	MZ
Social	Voter Turnout	INKAR
	Vote Share CDU	INKAR
	Vote Share SPD	INKAR
	Share Social Assistance	INKAR
	Mean ISEI	MZ
	Gini ISEI	MZ

Notes: This table displays all regional indicators considered for our analysis. The third column reports the data source, which is either the INKAR database (see description below), the Statistical Office of Germany (Destatis), the Hochschulrektorenkonferenz (HRK) or the Mikrozensus (MZ).

Table F.1 displays all 71 regional indicators we use as predictors in the Random Forest algorithm. In a first step, we retrieve data from Federal Institute for Building, Urban Affairs and Spatial Research (BBSR), which maintains the INKAR database of regional indicators (https://www.inkar.de/). These data are collected from various government bodies in Germany, including the German Statistical Office Destatis and the Institute for Employment Research (IAB). We select all indicators which think may potentially be relevant for mobility and are not too collinear to each other (for example, we do not include the general unemployment rate and the unemployment rates among males and females at the same time). In a second step, we add data from Destatis publications with information on the share of Gymnasium students among all secondary school students and compute the distance of the geographical center of each LLM to the next college based on data from the website of the Hochschulrektorenkonferenz (HRK; https://www.hochschulkompass.de/hochschulen/downloads.html). In a third step, we compute additional statistics on the LLM level in the MZ data, like the Gini coefficient in household income, the local A-Level wage premium or the ISEI (an international index of social status). We construct our final indicators as the time averages of variables over the years 2011 to 2018 at the LLM level.