Social Mobility in Germany*

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

March 2, 2022

Abstract

We characterize intergenerational mobility in Germany using census data on educational attainment and parental income for 526,000 children. Motivated by Germany's tracking system in secondary education, our measure of opportunity is the A-Level degree, a requirement for access to university. A 10 percentile increase in parental income rank is associated with a 5.2 percentage point increase in the A-Level share. This gradient remained unchanged for the birth cohorts 1980-1996, despite a large-scale expansion of upper secondary education. At the regional level, there exists substantial variation in mobility estimates. Local characteristics, rather than sorting patterns, 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:** University of St. Gallen (dominik.sachs@unisg.ch); **Schüle:** ifo Munich and LMU Munich (paul.schuele@econ.lmu.de). We thank Klaus Adam, Timm Bönke, 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 backgrounds are impeded from realizing their potential. Despite its importance, reliable mobility statistics are not available for many countries. Measuring social mobility is challenging, as it requires data that allow to link parental outcomes to a measure of opportunities for children. Household panel studies may contain this information but are typically too small to deliver sufficiently precise estimates for regional comparisons or the analysis of time trends (Lee and Solon, 2009; Mazumder, 2018). An attractive alternative are administrative data sources, such as linked tax records (e.g. Chetty et al., 2014). As in many other countries, however, such data is not available in Germany, where to date no large-scale empirical study of social mobility across time and space exists.

In order to fill this gap, this paper proposes and implements a new measurement strategy for social mobility in Germany. Motivated by Germany's early tracking system in secondary education, our mobility statistics measure the association between parental income and the educational opportunities of children. Our measure of opportunities captures whether a child will obtain the A-Level (Abitur), the highest secondary schooling degree in Germany. Since secondary school aged children and adolescents typically still live in the parental household, we are able to link them to their parents in the German census data. Our data covers one percent of the German population in every year from 1997 to 2018, providing detailed information on the educational activities of 526,000 children 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 parental income rank was associated with a 5.2 percentage point increase in the probability of obtaining an A-Level degree. For the birth cohorts 1980-1996, this parental income gradient has not changed despite a large-scale expansion of upper secondary education in Germany, the *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 the *Bildungsexpansion* took place uniformly across the income distribution, with almost identical increases in the share of A-Level educated children in all quintiles of the parental income distribution. This enhanced the odds ratio for disadvantaged children, but left the parental income gradient 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 geographic variation in social mobility across German states, cities, and local labor markets. For example, the top-bottom gap in the probability of obtaining an A-Level degree is 20 percentage points larger in Bremen than in Hamburg, two city states approximately 100 kilometers apart. We also find significant and meaningful differences within states. For example, the top-bottom gap is 8 percentage points larger in Cologne than in Düsseldorf, two large cities in North Rhine-Westphalia located approximately 40 kilometers apart. This is remarkable, as education policies, which prior literature has suspected to be a key determinant of mobility, vary mainly at the state level in Germany.

Third, we show that household characteristics can explain only a small fraction of the variation in mobility measures across local labor markets. This is important, as, abstracting from estimation uncertainty, differences in mobility estimates can arise either due to structural differences between places or due to systematic sorting of households into different local labor markets. Which answer prevails has important implications for the usefulness of place-based policies intended to promote social mobility, a topic of ongoing academic debate. The census data employed in this paper contains rich information on the structure and characteristics of households, allowing us to directly test the importance of sorting by conditioning on an extensive set of household characteristics. We find that the mobility ranking between local labor markets is largely unchanged when conditioning on household characteristics, indicating that sorting is unlikely to explain the observed regional variation.

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 association between income ranks of children and their parents using administrative tax data (Chetty et al., 2014; Dahl and DeLeire, 2008). While administrative tax returns offer key advantages like a large sample size and a high data quality in the upper half of the income distribution, this data comes with its own limitations and is 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 a child after secondary school is an appealing alternative for several reasons. Most importantly, it allows measuring mobility using census data, where children can only be linked to their parents as long as they live in the same household. As opposed to household surveys, census data offers a sample size which allows for regional comparisons or the analysis of time trends. As opposed to administrative tax data, census data is representative by construction and may contain detailed information on the socio-economic situation of households. Moreover, we can obtain mobility statistics for relatively recent cohorts, which is not possible when estimating income mobility.[1]

Due to its early-age tracking system, Germany is particularly suited for studying social mobility through the lens of educational opportunities. Only completion of the highest track grants the A-Level degree and thus 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 German society. More broadly, a large literature shows that educational attainment has intrinsic value and predicts a wide range of non-pecuniary outcomes (Lochner, 2011; Oreopoulos and Salvanes, 2011). Educational attainment as a measure of opportunity is thus a strong and comprehensive indicator for the opportunities of an individual in the German context. 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.

[1] A well established literature documents life cycle bias in income mobility estimates (Haider and Solon, 2006; Nybom and Stuhler, 2016; Solon, 1992). 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 measuring 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. Also in light of the increased demand for more "real-time" data in the wake of the COVID-19 pandemic, we consider the timeliness of our measure to be a key advantage over estimates of intergenerational income mobility.

The remainder of this paper is organized as follows. Section 2 discusses the related literature and relevant aspects of the German institutional framework. In Section 3, we describe data and measurement strategy. Section 4 reports our results at the national level. Regional estimates, including the analysis of local labor markets, are presented in Section 5. Section 6 concludes.

2 Related Literature and Institutional Background

2.1 Related Literature

The study of social mobility has a long tradition in economics, sociology and educational research. Across disciplines, efforts to understand and describe the association between the opportunities of children and their parents' socioeconomic status have been made. Since opportunities are difficult to measure, empirical studies of social mobility have generally aimed at the joint distribution of outcomes, with different disciplines emphasizing different outcomes. While early sociological studies focused on occupational transitions between generations, educational research studied intergenerational correlations in educational attainment.

The literature in economics, reviewed in Black and Devereux (2011), has traditionally measured social mobility by the intergenerational elasticity of (lifetime) earnings (IGE), which can be derived from standard intergenerational life-cycle models of human capital accumulation (e.g. Becker and Tomes, 1979; Mazumder, 2005; Solon, 1992). Since estimates of the IGE are sensitive to non-linearities and measurement issues at the bottom of the income distribution, recent empirical work relies on rank-rank correlations in lifetime income (Chetty et al., 2014; Dahl and DeLeire, 2008) to produce more robust mobility statistics. A major step forward in terms of data quality has been achieved by Chetty et al. (2014), who linked administrative tax records. The large sample size of this study has allowed the authors to produce reliable estimates of rank-rank correlations across regions in the US, opening up the field of research for a better understanding of the causes of social mobility (Chetty and Hendren, 2018). This approach was recently replicated for other countries, including Italy (Acciari et al., 2019), Switzerland (Chuard and Grassi, 2020), Canada (Corak, 2020) and Australia (Deutscher and Mazumder, 2020).

While measures of social mobility based on the joint distribution of lifetime incomes are attractive, as they allow for easy cross-country comparisons and have a natural interpretation in terms of consumption, they have important limitations. First, since they rely on estimates of children's lifetime income, they are only feasible for relatively old birth cohorts, as reliable estimates of lifetime income require data on children's earnings in the age range 30-40. These measures are therefore not suited for investigating recent developments in social mobility. Second, since individuals value non-monetary qualities of jobs (Kalleberg, 1977, 2011; Mottaz, 1985) and parental income is documented to be positively associated with the non-monetary compensation that children receive from their work (Boar and Lashkari, 2021), measures based on the joint distribution of incomes may overestimate the degree of intergenerational mobility. Finally, large-scale linked tax data is not available in many countries. In these countries, mobility measures based on the joint distribution of incomes can only be estimated with sufficient precision at the aggregate level, preventing further analysis of time and geographic variation that is possibly informative about the determinants of social mobility.

In Germany, it is not possible to link individual tax returns. For that reason, most empirical evidence on income mobility is based on the German Socio-Economic Panel (GSOEP), the German counterpart of the Panel Study of Income Dynamics (PSID). Like the PSID, the GSOEP provides detailed information about child outcomes and parental background, but suffers from a small sample size. In the GSOEP it is therefore not possible to document time trends or more fine-grained geographic variation in social mobility with a sufficient degree of statistical confidence. Schnitzlein (2016) shows that estimates of the national IGE 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 not coherent. 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 income mobility in Germany as compared to the US, and lower levels of mobility in East than in West Germany, albeit with high statistical uncertainty.

Our measurement approach circumvents the data requirements imposed by life-cycle bias concerns by focusing on children's educational opportunities, while retaining the interpretability advantages of income based measures of parental socioeconomic status. This allows us to draw on the German census data, providing us with the statistical power necessary to conduct a more comprehensive study of social mobility in Germany.[2] [3]

A similar approach was followed by Hilger (2015) for the US, who reports mobility statistics based on census data that measure the association between children's years of schooling and parental income. However, while we also rely on the co-residency of children and their parents, the outcome studied in Hilger (2015) manifests much later in life, when most children have already left the parental household, exacerbating sample selection concerns that necessitate an imputation procedure. Focusing on the years of schooling is necessary in the US context, where almost all children attend academic high school programs. In contrast, the German system of secondary education is separated in an academic and a vocational track, making it better suited for a census based analysis of social mobility as we outline below.

2.2 Institutional Background

The salient feature of Germany's system of secondary education is early age tracking, where only the successful completion of the highest track results in the award of an A-Level degree (Abitur) and grants direct access to the tuition-free national university system. After finishing the four-year[4] elementary school around the age of 10, children are allocated into one of three tracks. While the highest track, the Gymnasium (grades 5-12/13), provides general academic education that aims to prepare children for college, the lower two tracks (grades 5–9/10) provide vocational training with a focus

[2] A less comprehensive version of the German Census data has previously been used to document differences in the intergenerational correlation in educational attainment between East and West Germany (Klein et al., 2019; Riphahn and Trübswetter, 2013).

[3] The idea of relying on educational outcomes of children that can be measured early in life has recently been popularized in a small but growing literature on educational mobility in developing countries (Alesina et al., 2021; Asher et al., 2020; Muñoz, 2021).

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

on preparing students for an apprenticeship.[5] The specific design of the tracking system in secondary education can vary across the 16 federal states which bear the main responsibility for the education system. However, there exist only minor differences in state-provided financing. In addition, 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.

Since the early educational careers of children have important consequences for the choices available to them at later stages, and early track choices are heavily influenced by parental characteristics (Dustmann, 2004), the German institutional framework is particularly suited for studying social mobility through the lens of educational opportunities. The importance of track choices for social mobility is reinforced by the fact that almost all primary and secondary schools as well as universities are state-funded, mostly based on student headcounts, resulting in a comparatively large equality in the endowments and quality between different schools and universities.

Consequently, 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 German society. Overall, this illustrates that, for children in Germany, the A-Level degree is a compelling measure of their social and economic opportunities.

---

[5]The rigor of the tracking system is mediated by the possibility of switching tracks. In particular, it is common that talented students from the medium track switch to the general high track or attend a specialized high track after they finish their vocational degree when they are around 16 years old. 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 Microcensus (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.

3 Data and Measurement Strategy

Our analysis is based on data of the German Microcensus (Mikrozensus, hereafter MZ), a large-scale annual representative survey of the German population administered by the Federal Statistical Office of Germany. The MZ is comparable to, but more detailed than, the American Community Survey and constitutes the largest survey program of official statistics in Europe. The survey was first administered in West Germany in 1957 and includes East Germany since 1991. It contains individual level data on a wide range of topics, including family status and linkage within the household, citizenship, labor market participation, income as well as information on educational activities and attainment for all members of the sampled households.

The MZ has several features that make it particularly suited for our research question. First, it allows us to reliably match children to their parents as long as they are still registered at their parents' household. By law, it is compulsory for individuals living in Germany to register at their place of residence, and the sampled households are obliged to provide information on each person registered in their household. Second, it contains fine-grained geographic information and is sufficiently large to permit the estimation of mobility statistics for single cohorts and regions. Since its inception, the survey was continuously improved and its design and institutional embeddedness offer several advantages over comparable national surveys as we outline below.

Sampling Design. Each year, a randomly selected 1% sample of the population living in Germany is asked to participate in the survey. By law, participation is mandatory for members of the selected households, which remain in the survey for at most four subsequent years. The primary sampling units consist of clusters of neighboring buildings and all households belonging to a sampled cluster are interviewed. The unit non-response rate is approximately 3%.[7] Each year, one quarter of the initially sampled clusters are replaced by new clusters, resulting in partial overlap of sampling units. Appendix A contains additional information on the survey and sampling design of the MZ. The detailed nature of the questionnaire together with the low non-response

[7]The non-response rate is driven by households that could not be reached and residents in shared accommodations (Statistisches Bundesamt, 2018), which we exclude from our sample. The item non-response rate in our sample for the survey questions that we utilize is typically below 1%.

rate and the large sample size allow us to mitigate measurement and sample selection concerns often brought forward in the context of survey data.

3.1 Variable Definition

Measuring Opportunities of Children. Motivated by the importance of the A-Level degree for children's future educational and labor market opportunities in the German institutional framework, we measure opportunities by a binary variable Y_i that is equal to one if a child has obtained, or is on track to obtain, a degree that is equivalent to an A-Level, and zero otherwise. Specifically, our outcome variable is equal to one if (i) a child has obtained a degree that qualifies 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] In the following, we refer to this outcome as an A-Level degree and characterize intergenerational mobility in terms of the conditional probabilities of obtaining an A-Level degree for children of different parental backgrounds.

Our outcome definition takes into account three considerations. First, while the MZ survey is conducted on a rolling basis, A-Level degrees are typically awarded in the second quarter of the calendar 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 for around 40% of the graduating cohort in each survey year. Second, since the share of children failing the final examination in a given year is low[10], including upper stage students allows us to capture children that can reasonably be expected to obtain an A-Level degree but rotate out of the survey before they do so. Finally, including younger children disproportionately increases sample size, as younger children are more likely to live with their parents. Table 1 displays the share of children living with at least one parent by age of the child, calculated from our data.

[8]We classify educational qualifications as equivalent to an A-Level if they grant access to the tuition-free national university system. This includes *Allgemeine Hochschulreife (Abitur)*, *Fachgebundene Hochschulreife* and *Fachhochschulreife*.

[9]The MZ data contains information on the type of school and grade level attended by all sampled children. Our 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]The national average failure rate is approximately 3 percent on average for the years 2010-2020. For an overview of the share of children failing the final examination see https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html

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 the fraction of individuals which live in the same household as at least one of their parents in the MZ waves 1997 to 2018 by age at observation.

Virtually all children younger than 15 still co-reside with at least one parent. However, the share of children living with their parents is decreasing with child age, especially after the legal age of 18. While 92% of the 18 year olds are living with at least one of their parents, this fraction drops to 44% for individuals at the age of 23. In Section 3.3, we discuss how the co-residency and move-out patterns observed in the MZ data affect the interpretation of our results.

Measuring Parental Background. We measure parental background by a household's self-reported monthly net income, excluding the income of all dependent children. Our income measure covers all sources of income, including labor income, business profits and social security transfers. To account for differences in need and standard of living by household composition, we scale all household incomes by the modified OECD equivalence scale.[11] We then compute the households' percentile ranks in the sample distribution of equivalized household income,[12] and assign each child the rank of their respective household, which we refer to as the parental income rank R_i.

We emphasize that our aim is *not* to estimate some causal effect of parental income on children's educational attainment.[13] Instead, our measure intends to capture relative advantages in family circumstances of some children relative to others in a fashion that allows for the construction of robust and easy to interpret mobility statistics. To that end, parental income ranks are conceptually attractive, as the relevance of financial resources and costly enrichment activities for different aspects of child development

[11] Figure B.2 demonstrates that the choice of the scaling factor is not influential for our results at the aggregate level. However, the empirical conditional expectation function of our A-Level indicator can be better approximated linearly when computing ranks based on equivalized incomes.

[12] In Appendix A we provide information on the sample income distributions and details on the construction of the rank variable.

[13] Associations between parental income and children's educational attainment can be generated in numerous ways, including biological mechanisms (Black et al., 2020) as well as environmental factors, e.g. by shaping the beliefs about returns to education (Attanasio et al., 2020; Boneva and Rauh, 2021).

TABLE 2. Monthly Child-related Expenditures of Single Child Households

Category	Total	Education	Health	Food	Culture	Mobility	Other
Top Decile	1212	83	113	156	205	85	244
Bottom Decile	424	28	11	104	47	29	65
Ratio	2.85	2.96	10.27	1.5	4.36	2.93	3.75

Notes: This table reports estimates of the monthly child-related expenditures in Euro of dual parent, single child households in the top and bottom decile of the German national income distribution for different expenditure categories. The data is reported in the 2018 Income and Consumption Survey (EVS) of the Federal Statistical Agency (Statistisches Bundesamt, 2021).

is widely recognized and there exists empirical evidence of significant disparities in child-related expenditures across the income distribution in Germany. Table 2 reports estimates of monthly child-related expenditures in different categories based on data of the 2018 Income and Consumption Survey (EVS) for dual parent households with single children in the top and bottom decile of the national income distribution.The estimates reveal substantial gaps in monthly expenditures on child-enrichment activities in categories such as education, health as well as culture and leisure activities, suggesting that parental income ranks are a suitable measure of parental background for the construction of mobility statistics in Germany.

The continuous measure of household income provided in the MZ data that we use to compute parental income ranks is not asked for directly in the survey but imputed by the Statistical Office. The survey respondents report their personal income in 24 predefined bins. The Statistical Office then transforms the personal binned income into a continuous variable, essentially randomizing individuals uniformly within each bin. In a second step, these values are summed up to a continuous measure of household income. We discuss potential implications of this procedure for the external validity of our mobility statistics in Section 3.3.

3.2 Mobility Statistics

The central building block of all mobility statistics reported in this paper are estimates of the probability of children attaining an A-Level degree conditional on parental income rank $E[Y_i|R_i]$. Consequently, all estimands are descriptive, in the sense that uncertainty about our mobility statistics stems only from the fact that we do not observe the full population of Germany.

Following the recent literature, we define two sets of mobility statistics with the aim of distinguishing between two mobility concepts: absolute and relative mobility. While measures of absolute mobility are informative about the level of opportunities for disadvantaged children, relative mobility measures seek to capture differences in opportunities between children of disadvantaged backgrounds relative to those of more advantaged backgrounds.

Absolute Mobility. We measure absolute mobility by the probability of obtaining an A-Level degree for a child from the bottom quintile of the parental income distribution:

$$\text{Q1} = E(Y_i|R_i \leq 20). \tag{1}$$

We refer to this estimand as the Q1 measure and estimate it by its sample analogue $\overline{Q1}$. A high value of the $Q1$ measure implies high absolute mobility, as it indicates that a large share of disadvantaged children are eligible to enter the university system.

Relative Mobility. We define two measures of relative mobility, both concerned with the difference in opportunities between children from low and high-income families. A simple non-parametric measure of relative mobility is the Q5/Q1 ratio:

$$\text{Q5/Q1} = \frac{E(Y_i|R_i > 80)}{E(Y_i|R_i \leq 20)}, \tag{2}$$

which captures the odds ratio of obtaining an A-Level degree for children from the top quintile relative to those in the bottom quintile of the parental income distribution. A high value of the $Q5/Q1$ ratio implies low relative mobility. For example, a ratio of $Q5/Q1 = 2$ means 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. Analogous to the $Q1$ measure, we estimate the $Q5/Q1$ ratio by its sample analogue $\overline{Q5}/\overline{Q1}$.

Next to the $Q5/Q1$ ratio, we also estimate a parametric statistic of relative mobility. As demonstrated in the results section of this paper, the empirical conditional expectation function, $\widehat{E[Y_i|R_i]}$, of our outcome given the parental income rank is close to linear in various partitions of our data. As a consequence, we can use a parsimonious parametric model to characterize relative mobility. Formally, we do so by approximating the respective conditional expectation function (CEF) by its best linear predictor, which is defined as

$$\theta_{LP} = \arg\min_{\theta} E[(Y_i - Z_i'\theta)^2],$$

with $Z_i = (1, R_i)'$ and $\theta = (\alpha, \beta)$. In practice, we estimate the model parameters by running an OLS regression of our outcome indicator on the parental income rank variable.

If the CEF is linear, the slope coefficient β measures the gap in the probability of obtaining an A-Level between children at the top and the bottom of the income distribution. We refer to the slope coefficient as the *parental income gradient* and report estimates of $\beta \times 100$, which captures the gap in percentage points, for improved readability. While the Q5/Q1 ratio measures the relative outcome difference between children at the top and the bottom of the income distribution, the parental income gradient characterizes the absolute outcome difference and is therefore not sensitive to the baseline probability of obtaining an A-Level in the underlying population of interest.

3.3 Sample Definition and Limitations

We use the MZ survey waves from 1997 to 2018, for which a consistent definition of all relevant variables is available.[14] Our primary sample contains all children aged 17-21 which are observed in the same single-family household as at least one of their parents. The age range is chosen to balance the following trade-off: For older children, our outcome is measured more precisely, i.e. we do not need to rely on upper-stage enrollment but are more likely to observe the completed degree. At the same time, the fraction of children in our sample that has already moved out of the parental household, and thus can not be matched to their parents, increases with age, which guides our choice for the upper bound. The lower bound is chosen as children enrolled in the upper stage of an A-Level track are typically at least 17 years old. In the following, we discuss potential concerns regarding the external validity of our mobility estimates.

[14]For our national and regional estimates, we restrict our sample to the survey waves 2011-2018 (231,000 children) to produce recent mobility statistics and avoid ambiguities caused by a series of administrative reforms that changed county boundaries. The mobility statistics by birth cohort reported in Section 4.2 are computed based on the 1980-1996 birth cohorts (526,000 children).

Sample Selection. An immediate concern caused by the observed move-out patterns in the MZ data relates to the representativeness of our sample. If the observed move-out decisions were systematically related to both parental income and the educational attainment of children, the external validity of our estimates would be undermined as our statistics would not measure social mobility in the population of interest. While we acknowledge that dependencies of this type are generally plausible, we do not find evidence of sample selection in our data. Table 3 documents how time-constant characteristics of the children in our sample change with the age at observation. If move-out

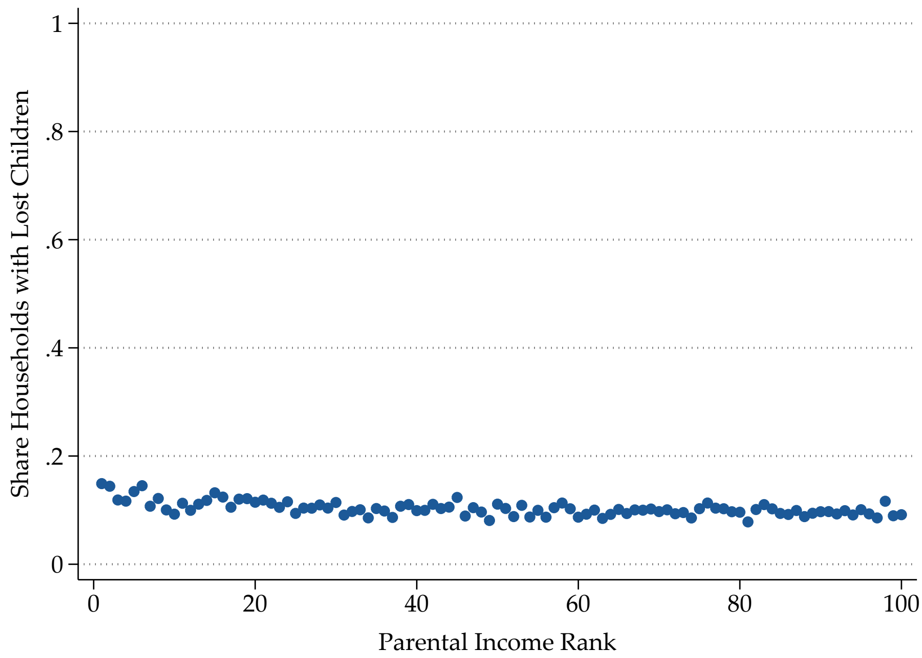
TABLE 3. Average Characteristics of Children by Age of Observation

Child Age	Share Female	Mean Parental Inc. (Equiv.)	Parental Inc. Rank	Share Parents with A-Level
17	0.49	1367	50	0.33
18	0.48	1367	50	0.32
19	0.47	1367	50	0.32
20	0.44	1359	50	0.31
21	0.42	1360	50	0.31

Notes: This table reports average attributes of children in the MZ waves 1997 to 2018 that are observed in the same household as at least one of their parents by age of observation. The ranks are computed based upon the sample distribution of equivalized household income as described in Section 3.3.

were to occur randomly, we should not see systematic changes in these statistics for older children for which the co-residency rate is lower. While move-out varies with social characteristics like gender, the average parental income and the associated income rank of children in the age range 17-21 are essentially constant. In addition, we can exploit the partial panel dimension of the MZ to investigate selection patterns more directly. Figure 1 displays the share of observed move-outs of children by parental income rank for the subsample of households in our data that is observed in the survey in multiple years. It shows that move-outs occur uniformly across the income distribution and are thus uncorrelated with parental income rank. Both exercises suggest that sample selection is not a major concern for our analysis. In addition, we demonstrate in the next section that choosing alternative age ranges barely affects our results.

FIGURE 1. Move-out Frequency by Parental Income Rank



Notes: This figure shows the relative frequency of move-outs of children aged 17-20 by parental income rank. It is computed based on a sample of 265,229 children in the years 2012-2018 where we observe the partial panel dimension of the MZ and can identify households surveyed for more than one wave. We define households with "lost children" as households which report a lower number of children aged 17-20 than in the previous year.

Measurement Error. A second concern relates to measurement error in parental income due to the binned nature of our income data. As explained in Section 3.1, any measurement error induced by the imputation procedure of the Statistical Office is independent of household characteristics. To the extent that our estimates suffer from measurement error, we should therefore expect attenuation bias in our mobility statistics. However, this type of measurement error is unlikely to be a concern for the mobility statistics under consideration: First, we note that the random imputation procedure of the Statistical Office already dampens measurement error by constructing household income as the sum of all personal incomes within the household. In addition, our measurement approach relies on rank-based measures of social mobility, documented to have favorable bias properties in the presence of measurement error relative to other approaches (Nybom and Stuhler, 2017).

Measurement error in parental income could further arise from transitory income shocks. As noted above, our definition of parental income does not seek to capture lifetime income but the parental resources available for enrichment activities during childhood. In our baseline estimates, we therefore compute parental income ranks based on yearly incomes. To the extent that parental incomes fluctuate from year to year due to transitory income shocks, these statistics could nevertheless overstate mobility. To address this issue, we again exploit the panel dimension of our data for the subsample of households that is observed in multiple survey waves and compute multi-year averages of parental income before assigning ranks. We demonstrate below that our results are insensitive to this procedure.

While such fluctuations could affect our rank computations, our education-based measure of opportunities does not suffer from life-cycle bias. In contrast to traditional measures that rely on the labor market incomes of children, we can therefore study recent birth cohorts without compromising the quality of our estimates.

Standard Errors. The standard errors reported alongside our estimates in the results section of this paper abstract from the fact that we estimate the cutoffs defining the percentile ranks. For the parental income gradient as well as the Q1 and Q5 measure, we report Liang-Zeger standard errors (Liang and Zeger, 1986) clustered at the level of the sampling district, the primary sampling unit of the MZ.[15] For the Q5/Q1 ratio, we report plug-in standard errors based on a "delta method argument", that is we linearize the ratio of averages which yields the following approximation for the variance of the sampling distribution of the $\overline{Q5}/\overline{Q1}$ sample ratio:

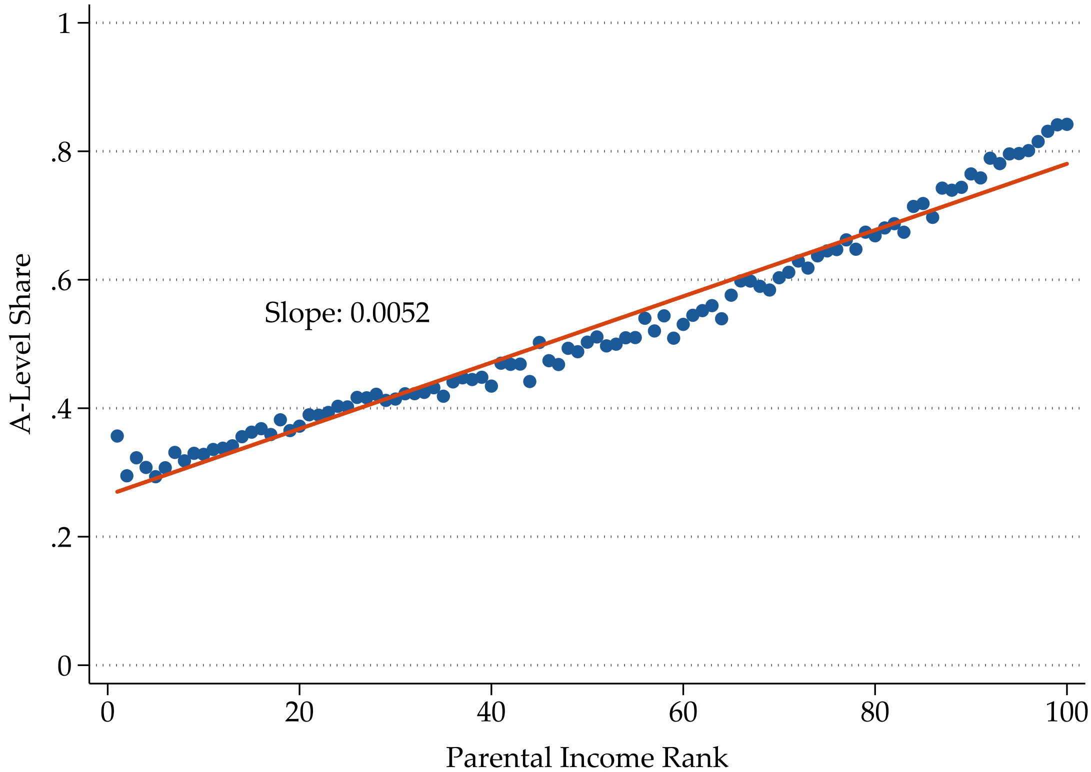
$$V(\overline{Q5}/\overline{Q1}) \approx \frac{1}{(Q1)^2}\left(V(\overline{Q5}) + \left[\frac{Q5}{Q1}\right]^2 V(\overline{Q1}) - 2\frac{Q5}{Q1}Cov(\overline{Q5},\overline{Q1})\right).$$

4 National Estimates

We begin our empirical analysis by characterizing social mobility at the national level. Figure 2 shows the share of children with an A-Level degree by parental income rank in our data, as well as the best linear approximation to the empirical CEF. As can be seen, a linear model provides a reasonable approximation to the CEF, a regularity that we observe in essentially all considered partitions of our data. In the national data,

[15]The MZ data allows for consistent identification of primary sampling units across waves following the 2011 survey. For the estimates in Section 4.2, where we also use prior waves, we instead cluster standard errors at the household level.

FIGURE 2. Social Mobility at the National Level



Notes: This figure shows the fraction of children aged 17-21 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree by percentile rank of their parents in the national income distribution for the MZ waves 2011-2018. The income ranks are computed with respect to the national income distribution of households with children aged 17-21 in each survey year. The reported slope coefficient of 0.0052 (SE 0.004) is estimated by OLS using the underlying micro data.

we estimate the parental income gradient at $\beta \times 100 = 0.52$, implying a gap of roughly 50% in the probability of obtaining an A-Level degree between children from the top and the bottom of the income distribution.[16] Our measure of absolute mobility in the national data suggests that one third of children from the bottom quintile of the income distribution complete an A-Level degree, with Q1 estimated at 0.34. Both parametric and non-parametric mobility statistics imply that the odds ratio in the probability of obtaining an A-Level degree between children from the top quintile relative to the bottom quintile is greater than 2, with Q5/Q1 estimated at 2.25. Consistent with the discussion of sample selection concerns in Section 3.3, we find that the estimates are

[16]For the national estimates, we pool our data over the period 2011-2018 to ensure consistency with the regional estimates in Section 5, for which obtaining results before 2011 is difficult due to frequent reforms of local administrative boundaries.

robust to the use of multi-year averages of parental income and to variations in the age restriction defining our sample, as shown in Table 4.

TABLE 4. National Estimates for Different Age-Restrictions

Age	Gradient	Q1	Q5	Q5/Q1	A-Level Share	N
17-21	0.52 (0.004)	0.34 (0.003)	0.76 (0.003)	2.25 (0.021)	0.52	230,972
17-21 (Averaged)	0.52 (0.004)	0.34 (0.003)	0.77 (0.003)	2.26 (0.022)	0.52	230,972
17	0.53 (0.007)	0.30 (0.005)	0.73 (0.004)	2.46 (0.042)	0.49	53,324
18	0.51 (0.007)	0.35 (0.005)	0.77 (0.004)	2.18 (0.033)	0.54	51,278
19	0.51 (0.008)	0.35 (0.005)	0.77 (0.005)	2.19 (0.035)	0.53	46,747
20	0.51 (0.008)	0.35 (0.005)	0.77 (0.005)	2.19 (0.036)	0.53	42,396
21	0.52 (0.008)	0.34 (0.006)	0.77 (0.005)	2.24 (0.039)	0.52	37,227

Notes: This table reports national mobility statistics for the MZ waves 2011-2018. The first row corresponds to our primary sample. The second row replicates these estimates using multi-year averages of parental income before assigning ranks as described in Section 3.3. The additional rows report estimates for samples containing only children of a given age at measurement, as indicated in the first column. The standard errors in parentheses are computed as described in Section 3.3.

Do these estimates depict Germany as a country of high or low relative mobility? While a cross-country comparison of our results is not straightforward, as the German system of upper secondary education and university funding is unusual, we are aware of two US studies which report comparable mobility statistics. Using data from the Census 2000, Hilger (2015) reports a parental income rank gradient of 3.6 percentage points in attending college for children aged 19-21. A higher point estimate is reported in Chetty et al. (2014), who estimate the rank gradient in college enrollment at 6.7 percentage points for children aged 18-21 based on tax registry data. Under the assumption that college enrollment conditional on having obtained an A-Level degree is weakly increasing in parental income rank, our estimate of 5.2 percentage points implies a college enrollment gradient that falls into the range of point estimates reported for the US. Abstracting from differences in the distributions of college quality and the selection of students of different parental backgrounds into colleges of different quality, our estimates suggest that educational mobility in Germany is similar to the US. We consider this finding noteworthy, as (after tax) income inequality is more pronounced in the US than in Germany, suggesting that one should expect steeper rank gradients in the US.[17]

Our finding contrasts with cross-country comparisons in relative income-mobility, which typically report higher mobility in Germany, highlighting the conceptual difference between income and education based measures of social mobility. Similar results were obtained by Landersø and Heckman (2017), who find that Denmark, a society that is characterized by high levels of income mobility, is similar to the US in terms of measures of educational social mobility.

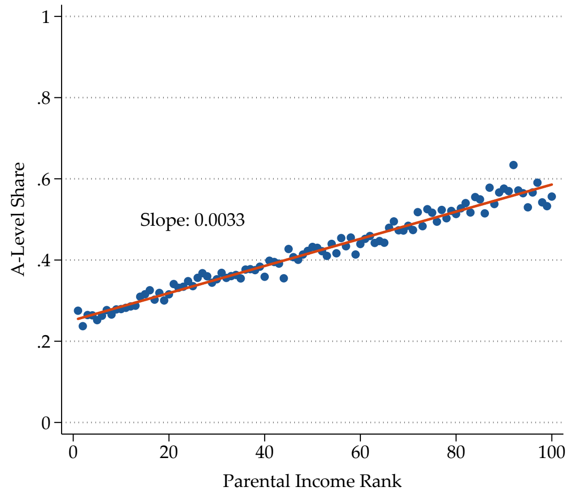
4.1 Subgroup Estimates

A natural question to ask is whether the national estimates mask meaningful differences in mobility measures across subpopulations. Table 5 reports mobility statistics for selected subgroups typically emphasized in the analysis of social mobility.

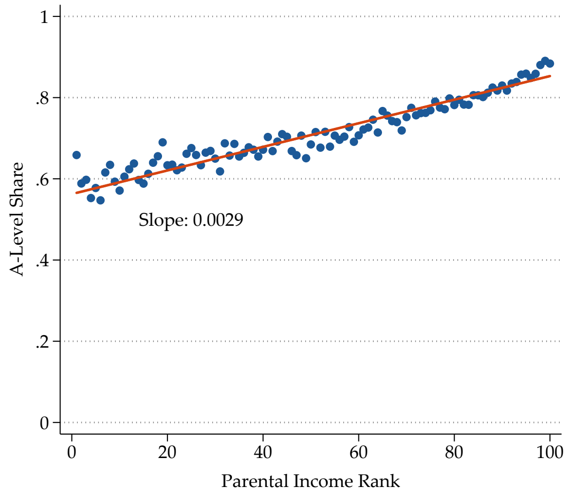
We document several interesting patterns. Most importantly, we find substantial differences by parental education. Figure 3 displays the A-Level share of children by parental income rank and the associated parental income gradient separately for children from households where no parent has an A-Level degree (Panel A) and for children from households where at least one parent has an A-Level degree (Panel B). The A-Level share among children of parents without an A-Level degree at the top of the income distribution is comparable to the A-Level share among children with at least one A-Level educated parent at the bottom of the income distribution. Roughly speaking, the empirical distribution for children of A-Level educated parents is shifted upwards by approximately 30 percentage points, uniformly across ranks. The conditional rank gradients are attenuated due to the positive correlation between parental

[17] Rauh (2017), for example, finds a negative correlation between inequality and public education expenditures across countries. If public education expenditures benefit lower-income children more, one expects a steeper rank gradient in the US. Remarkably, our results show this is not the case.

FIGURE 3. Differences by Parental Education



(A) No Parent with A-Level



(B) At Least One Parent with A-Level

Notes: This figure shows the fraction of children aged 17-21 observed in the MZ survey waves 2011-2018 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree by parental income rank, separately for children of parents who have not obtained an A-Level degree (Panel A) and children of parents where at least one of the parents has obtained an A-Level degree (Panel B). The ranks are computed based upon the sample distribution of equivalized household income as described in Section 3.3. The reported estimates of the parental income gradient are based on the underlying micro data. Standard errors are reported in the first panel of Table 5.

education and income ranks, with point estimates of approximately 0.3 in both groups. The intergenerational correlation in A-Level attainment in our data is 0.54. This finding highlights that the interpretability advantages of income-only based measures of parental background come at the cost of missing observable attributes of households that could be used to characterize social mobility more comprehensively.

The estimates reported in Table 5 reveal a few more interesting discrepancies. At the bottom of the income distribution, females and children with migration background are approximately 11 and 4 percentage points more likely to obtain an A-Level degree than their respective male and native counterparts. While the gender-gap is close to constant across the income distribution, the difference between migrant and native children vanishes in the top quintile. Moreover, we document larger income rank gradients for children of married and cohabiting couples, as well as for natives and children living in East Germany. The East-West gap in parental income gradients is 0.1, implying a 10 percentage points larger top-bottom gap in the probability of attaining an A-Level degree in East Germany as compared to West Germany.

TABLE 5. Mobility Statistics for Subgroups

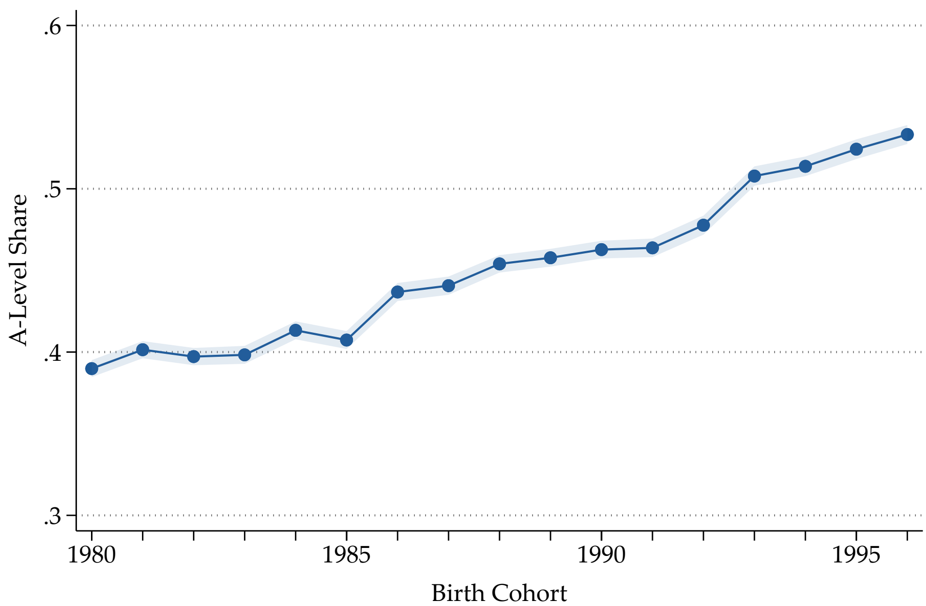
		Gradient	Q1	Q5	Q5/Q1	A-Level	N
Parental Education	No A-Level	0.33 (0.006)	0.28 (0.003)	0.55 (0.006)	1.94 (0.028)	0.39	145,892
	A-Level	0.29 (0.007)	0.61 (0.007)	0.84 (0.003)	1.36 (0.016)	0.75	85,080
Parenting Status	Single Parent	0.50 (0.010)	0.34 (0.004)	0.72 (0.009)	2.13 (0.037)	0.47	50,622
	Two Parents	0.54 (0.005)	0.34 (0.004)	0.76 (0.003)	2.26 (0.027)	0.54	179,715
Parents Married	Not Married	0.46 (0.010)	0.33 (0.004)	0.69 (0.008)	2.12 (0.037)	0.47	51,018
	Married	0.54 (0.005)	0.35 (0.004)	0.77 (0.003)	2.22 (0.025)	0.54	172,999
Gender	Male	0.53 (0.006)	0.29 (0.004)	0.72 (0.004)	2.49 (0.033)	0.47	123,649
	Female	0.50 (0.006)	0.40 (0.004)	0.81 (0.003)	2.02 (0.023)	0.58	107,323
Migration Status	Native	0.55 (0.005)	0.32 (0.004)	0.76 (0.003)	2.35 (0.028)	0.54	164,018
	Migrant	0.47 (0.009)	0.36 (0.004)	0.75 (0.007)	2.11 (0.032)	0.48	60,908
Region	West	0.50 (0.005)	0.34 (0.003)	0.76 (0.003)	2.19 (0.022)	0.52	201,684
	East	0.60 (0.011)	0.31 (0.007)	0.80 (0.007)	2.61 (0.062)	0.51	29,288
Siblings	Yes	0.55 (0.005)	0.35 (0.003)	0.79 (0.003)	2.29 (0.024)	0.52	156,960
	No	0.49 (0.007)	0.32 (0.005)	0.72 (0.004)	2.27 (0.039)	0.52	74,012
Birth Order	1st Child	0.51 (0.005)	0.34 (0.003)	0.76 (0.003)	2.22 (0.023)	0.53	165,336
	2nd Child	0.52 (0.008)	0.34 (0.005)	0.77 (0.005)	2.27 (0.036)	0.51	56,996
	Later Child	0.57 (0.021)	0.31 (0.009)	0.78 (0.017)	2.48 (0.092)	0.45	8,640

Notes: This table reports mobility statistics for selected groups of children observed in the MZ survey 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. The standard errors reported in parentheses below each point estimate are computed as described in Section 3.3.

4.2 Time Trends

We next ask how social mobility has evolved over time. While our descriptive approach does not allow us to attribute changes in mobility measures to specific policies, our measurement strategy enables us to provide novel evidence on the evolution of social mobility in Germany for relatively recent birth cohorts. The period we study is particularly interesting, as it covers the second half of the arguably most significant educational reform in post-war Germany, the *Bildungsexpansion*, a large-scale policy of expanding upper secondary and higher education that, starting in the early 1970s, increased the A-Level share from around 20% to approximately 50% for the birth cohorts since the mid 1990s. This expansion was a policy response to a heated public debate on social mobility (Dahrendorf, 1966) and the increasing importance of education for economic growth at the time (Hadjar and Becker, 2006; Picht, 1964). We ask whether the large-scale expansion of upper-secondary education in Germany was accompanied by changes in social mobility as defined by our mobility measures.

FIGURE 4. A-Level Share by Cohort



Notes: This figure shows the fraction of children born between 1980 and 1996 and observed at ages 17-21 that are either enrolled in the upper stage of an A-Level track or attained an A-Level degree in the MZ data. The shaded area displays pointwise 95% confidence intervals based on standard errors as described in Section 3.3.

To this end, we focus on a sample of 526,000 children born between 1980-1996.[18] At the time of writing, the children of the respective birth cohorts are 25-40 years old and constitute a significant part of the German working population. Including relatively young cohorts in our analysis is feasible, as, in contrast to traditional measures that rely on the labor market incomes of children, our education-based measure of opportunities does not suffer from life-cycle biases. Figure 4 depicts the evolution of the A-Level share among 17-21 year old children in the MZ data for the birth cohorts under consideration. Our data covers roughly the second half of the expansion, with an observed increase in the A-Level share of 14 percentage points from 39% for the 1980 birth cohort to 53% for children born in 1996.[19] At the same time, income inequality increased only moderately,[20] and we do not find evidence that the expansion was accompanied by a decline in the A-Level wage premium, as documented in Appendix Figure B.1.

Figures 5 and 6 display estimates of our mobility measures for the same cohorts. While the odds ratio captured by the Q5/Q1 ratio decreased by approximately one third, from around 3 for the 1980 birth cohort to slightly above 2 for the 1996 cohort, the parental income gradient has remained constant at around 0.52, the point estimate that we report at the national level based on more recent data. At the same time, absolute mobility as measured by the Q1 measure increased substantially, from approximately 0.22 in 1980 to 0.35 in 1996. The same overall pattern emerges when estimating mobility trends for the subgroups studied in Section 4.1 as reported in Figures B.4 and B.5 in the Appendix.

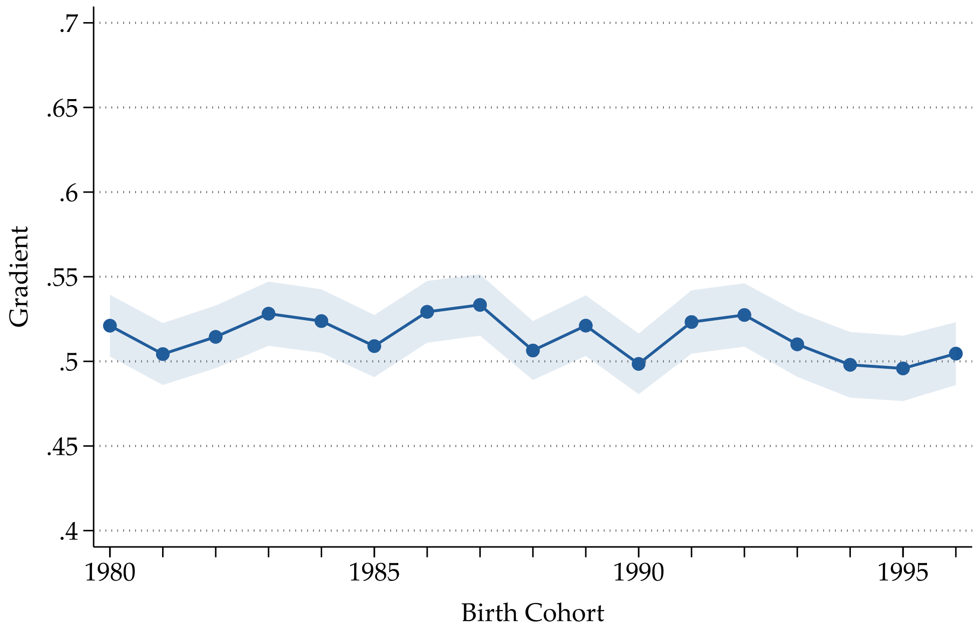
The connection between these findings is best summarized in Figure 7, which depicts the A-Level share by quintile across birth cohorts: The *Bildunsexpansion* took place uniformly across the income distribution, with increases of about 14 percentage points

[18] We restrict our attention to these cohorts to rule out that our estimates are affected by 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.

[19] The *Bildungsexpansion* featured a parallel increase of tertiary education and did not decrease the share of A-Level graduates taking up university studies. In the years 2002-2015, where most of our birth cohorts graduate, it fluctuated around 70% (https://www.datenportalbmbf.de/portal/de/Tabelle-2.5.74.html).

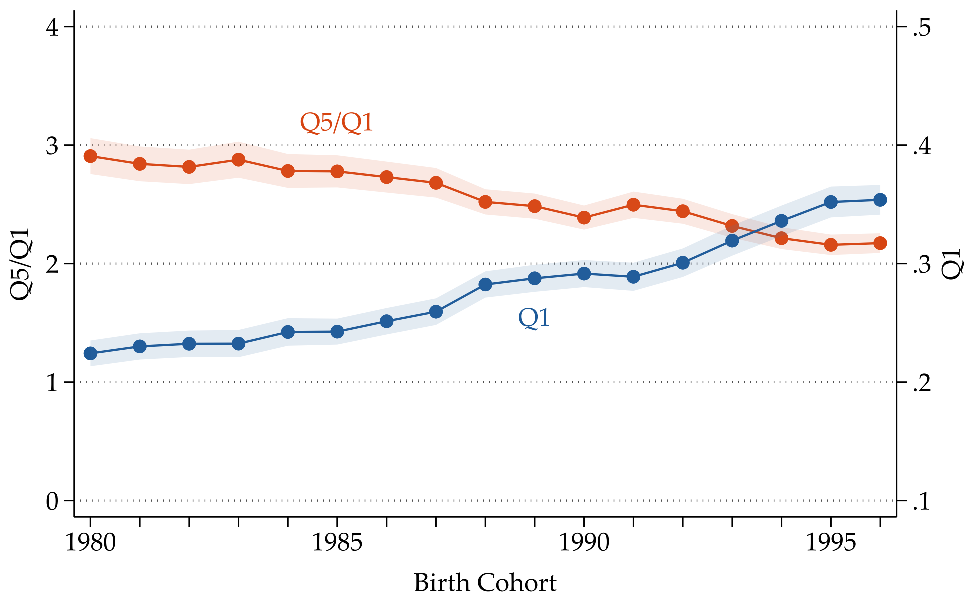
[20] While wage inequality rose in the 1990s and early 2000s when most children in our sample grew up, Fuchs-Schündeln et al. (2010) document that inequality in consumption and disposable income, the income concept used in this paper, increased only moderately.

FIGURE 5. Parental Income Gradient by Cohort



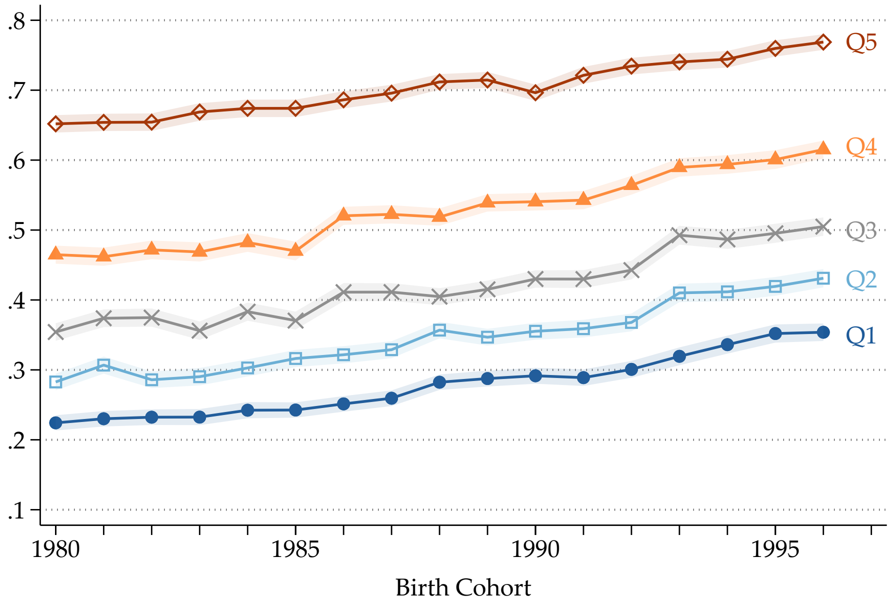
Notes: This figure shows for children aged 17-21 the evolution of the parental income gradient by birth cohort. The shaded area displays pointwise 95% confidence intervals based on standard errors as described in Section 3.3.

FIGURE 6. Quintile Measures by Cohort



Notes: This figure shows for children aged 17-21 the evolution of the quintile based measures of social mobility by birth cohort. While the left axis corresponds to the Q5/Q1 ratio, the right axis corresponds to the Q1 measure. The shaded areas display pointwise 95% confidence intervals based on standard errors as described in Section 3.3.

FIGURE 7. A-Level Share by Cohort Quintile



Notes: This figure shows the share of children born between 1980 and 1996 who obtained an A-Level degree by birth cohort and quintile of the parental income distribution in the MZ data. The shaded area displays pointwise 95% confidence intervals based on standard errors as described in Section 3.3.

in the A-Level share in all parts of the distribution. Did the *Bildunsexpansion* achieve its goal of fostering social mobility in Germany? While the expansion unquestionably increased absolute mobility as we measure it, the time trend in relative mobility is less straightforward to interpret. On the one hand, the attenuation of the Q5/Q1 ratio caused by the uniform increases in A-Level shares could suggest an increase in relative mobility according to a proportional notion of the concept. On the other hand, a less optimistic angle to interpret the same development is to consider the inverse odds ratio, that is the ratio between the probability *not* to obtain an A-Level for children in both quintiles. In the birth cohort 1980, children in Q1 were 2.2 times more likely not to obtain an A-Level degree than 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 has actually widened. In contrast, the unaltered top-bottom gap in the probability of attaining an A-Level captured by the parental income gradient emphasizes stagnation in absolute differences. As the parental income gradient is insensitive to the chosen reference point, we tend to interpret the evidence primarily as a stagnation of relative mobility. However, as both absolute and relative disparities often form the normative basis for interventions, all readings can be justified.

5 Regional Estimates

An interesting regularity documented in the recent empirical literature on social mobility is that there exists substantial geographic variation in social mobility measures within politically homogeneous entities, suggesting that regional comparisons can be used to gain a better understanding of the causes of social mobility (e.g. Acciari et al., 2019; Chetty et al., 2014; Chuard and Grassi, 2020; Corak, 2020; Deutscher and Mazumder, 2020). This idea is appealing, as attributing cross-country discrepancies in social mobility to differences in single characteristics or policies is difficult to justify. Complementary to well-designed evaluations of political reforms that rely on variation across time (e.g. Bertrand et al., 2021), within-country geographic variation can be helpful in understanding the causal mechanisms fostering or impeding social mobility by identifying exposure effects (Bütikofer and Peri, 2021; Chetty and Hendren, 2018). Moreover, pronounced regional differences can suggest mechanisms that warrant investigation.

The regional analysis conducted in this section is motivated by these considerations. In a first step, we present evidence of meaningful geographic variation in our mobility measures across regions in Germany. In a second step, we then ask what we can learn from the observed differences. We structure our regional analysis by disaggregating our data in a stepwise fashion, lending credence to our parametric mobility statistics while taking into account the political and economic landscape of Germany.

5.1 States

A natural starting point for our regional analysis are the 16 federal states of Germany. By constitutional law, the responsibility for the design and implementation of the education system falls under the jurisdiction of the German states and not under the jurisdiction of the federal government. As a consequence, the states have considerable

TABLE 6. Social Mobility at the State Level

State	Gradient	Q1	Q5	Q5/Q1	A-Level Share	Tracks	Binding Rec.
Hamburg	0.45 (0.033)	0.43 (0.023)	0.80 (0.017)	1.86 (0.109)	0.60	2	No
Rhineland-Palatinate	0.50 (0.019)	0.36 (0.013)	0.76 (0.011)	2.12 (0.086)	0.53	2	No
North Rhine-Westphalia	0.51 (0.009)	0.41 (0.006)	0.82 (0.005)	2.02 (0.032)	0.59	3	Ref
Hesse	0.52 (0.015)	0.39 (0.011)	0.81 (0.007)	2.07 (0.061)	0.59	3	Ref
Baden-Württemberg	0.52 (0.011)	0.34 (0.008)	0.76 (0.006)	2.24 (0.056)	0.53	3	Ref
Saarland	0.53 (0.040)	0.33 (0.024)	0.74 (0.025)	2.28 (0.186)	0.54	2	Ref
Schleswig-Holstein	0.53 (0.023)	0.32 (0.015)	0.76 (0.014)	2.34 (0.117)	0.52	2	No
Lower Saxony	0.54 (0.013)	0.29 (0.008)	0.73 (0.009)	2.52 (0.077)	0.48	3	Ref
Bavaria	0.54 (0.011)	0.24 (0.007)	0.67 (0.006)	2.75 (0.084)	0.42	3	Yes
Berlin	0.56 (0.021)	0.39 (0.013)	0.85 (0.011)	2.20 (0.082)	0.59	2	No
Brandenburg	0.57 (0.027)	0.35 (0.019)	0.84 (0.014)	2.37 (0.134)	0.60	2	Ref
Saxony-Anhalt	0.58 (0.034)	0.25 (0.017)	0.72 (0.026)	2.88 (0.227)	0.43	2	Ref
Saxony	0.61 (0.025)	0.28 (0.014)	0.78 (0.016)	2.83 (0.156)	0.48	2	Yes
Mecklenburg-Vorpommern	0.63 (0.041)	0.25 (0.020)	0.76 (0.028)	3.00 (0.256)	0.45	2	No
Bremen	0.64 (0.044)	0.32 (0.025)	0.86 (0.026)	2.65 (0.220)	0.55	2	No
Thuringia	0.65 (0.032)	0.25 (0.017)	0.76 (0.023)	3.07 (0.234)	0.46	2	Yes

Notes: This table reports mobility statistics for each federal state of Germany based on all children observed in the MZ waves 2011-2018. The standard errors reported in parentheses below each point estimate are computed as described in Section 3.3. The classification of the state education systems is based on the description of educational reforms in Helbig and Nikolai (2015). In the last column, "Ref" indicates that teacher recommendations were reformed during the time period relevant for our analysis.

discretion in the design of their education systems, leading to distinctions in the rigor of the tracking system, the capacities of each track, the types of schools and curricula and other important features of the education system.

In particular, states differ with respect to the duration of primary school after which all children are allocated into the different tracks, the number of tracks (2 or 3) and the importance of teacher recommendations for admitted track choices. While in all states teachers recommend a track for each child at the end of primary school, track recommendations are binding only in some states. These parameters of the state education systems and their suspected consequences for social mobility are often at the center of the public debate on educational mobility in Germany.

Table 6 reports our mobility estimates for the 16 states, sorted by the point estimate of the parental income gradient in ascending order. We document significant and economically meaningful differences in both absolute and relative mobility measures between states. For example, the top-bottom gap in the probability of attaining an A-Level degree is approximately 20 percentage points larger in Bremen than in Hamburg, two city states in north-west Germany approximately 100 kilometers apart. Similarly, the share of children obtaining an A-Level degree from the bottom quintile of the parental income distribution is 10 percentage points larger in Baden-Württemberg than in Bavaria, the two southernmost states of Germany. The estimated differences between states do not result from differences in the shape of the empirical CEFs, as we find that the linearity assumption underlying our parametric mobility estimates is supported by the data (compare Figure B.6). The table also reiterates the east-west gap documented in Section 4.1: except for Bremen, the least mobile states are all located in East Germany.

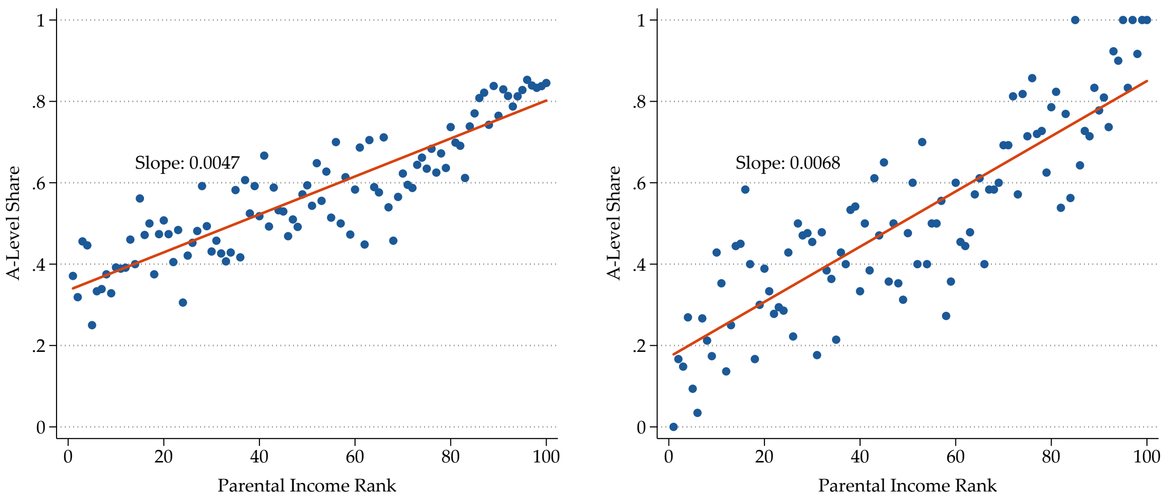
While we find that the differences in our measure of absolute mobility can be well explained by differences in the states' A-Level shares, that is the relative capacity of the highest track, there is no clear pattern in our estimates with respect to the aforementioned characteristics of the state education systems displayed in the last two columns of the table. Our findings suggest that, while certainly important, the design of the tracking system cannot readily explain the pronounced differences in our mobility measures between states.

5.2 Cities

A similar picture emerges when we restrict our analysis to urban regions of Germany. Table 7 reports our mobility estimates for the 15 largest labor markets of Germany, consisting of cities and their catchment areas as defined by commuting flows.

Compared to the national average, the largest urban regions of Germany show lower levels of relative, but higher levels of absolute social mobility. At the same time, the table shows that the regional differences observed at the state-level can also be found within states. For example, the top-bottom gap is approximately 8 percentage points larger in Cologne than in Düsseldorf, two large cities in North Rhine-Westphalia located approximately 40 kilometers apart. Similarly, our estimates of absolute mobility differ by 8 percentage points between Nuremberg and Munich, two large cities in Bavaria. The most striking discrepancy between cities in our data is observed for Hamburg and Leipzig, with a difference of approximately 20 percentage points in the estimated top-bottom gap, as well as 15 percentage points in our estimate of the Q1 measure. Figure 8 displays our raw data for the two cities.

FIGURE 8. Social Mobility in Hamburg and Leipzig



(A) Hamburg (B) Leipzig

Notes: This figure shows the fraction of children aged 17-21 observed in the MZ survey waves 2011-2018 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree in Hamburg (Panel A) and Leipzig (Panel B). The reported slope coefficients are estimated by OLS using the underlying micro data. Standard errors are reported in Table 7.

TABLE 7. Social Mobility in the 15 Largest Urban Labor Markets

City	State	Gradient	Q1	Q5	Q5/Q1	A-Level Share
Hamburg	HH/SH	0.47 (0.025)	0.41 (0.018)	0.79 (0.012)	1.94 (0.090)	0.58
Düsseldorf	NW	0.47 (0.029)	0.45 (0.023)	0.84 (0.014)	1.87 (0.100)	0.65
Münster	NW	0.47 (0.041)	0.47 (0.030)	0.84 (0.021)	1.78 (0.120)	0.62
Gelsenkirchen	NW	0.50 (0.035)	0.40 (0.018)	0.81 (0.029)	2.01 (0.116)	0.57
Stuttgart	BW	0.50 (0.024)	0.34 (0.017)	0.75 (0.012)	2.19 (0.114)	0.55
Bonn	NW	0.50 (0.039)	0.44 (0.030)	0.86 (0.016)	1.94 (0.135)	0.65
Duisburg	NW	0.51 (0.033)	0.42 (0.022)	0.84 (0.017)	2.02 (0.113)	0.58
Frankfurt	HE	0.52 (0.025)	0.42 (0.019)	0.83 (0.011)	1.97 (0.093)	0.62
Munich	BY	0.54 (0.025)	0.31 (0.021)	0.71 (0.011)	2.32 (0.162)	0.53
Dortmund	NW	0.55 (0.033)	0.40 (0.022)	0.86 (0.017)	2.16 (0.125)	0.59
Cologne	NW	0.55 (0.027)	0.38 (0.019)	0.85 (0.014)	2.25 (0.120)	0.60
Hanover	NI	0.56 (0.036)	0.30 (0.022)	0.76 (0.021)	2.51 (0.195)	0.53
Berlin	BE	0.56 (0.021)	0.39 (0.013)	0.85 (0.011)	2.20 (0.082)	0.59
Nuremberg	BY	0.60 (0.035)	0.23 (0.022)	0.70 (0.023)	3.01 (0.297)	0.43
Leipzig	SN	0.68 (0.044)	0.26 (0.026)	0.80 (0.028)	3.11 (0.335)	0.48

Notes: This table reports mobility statistics for the 15 largest urban local labor markets in Germany, as measured by their total population in 2017, based on the MZ waves 2011-2018. The local labor markets are sorted, in ascending order, by the point estimate of the parental income gradient. Standard errors are computed as described in Section 3.3. The point estimates for the city-states can differ from those reported in Table 6, as the urban labor markets typically also include surrounding towns and villages.

Similar to the previously considered partitions of our data, we show in Figure B.7 that the empirical CEFs are well approximated by a linear function. Overall, our city-level findings suggest that the relative opportunities of children can differ meaningfully across politically similar and geographically close regions of Germany.

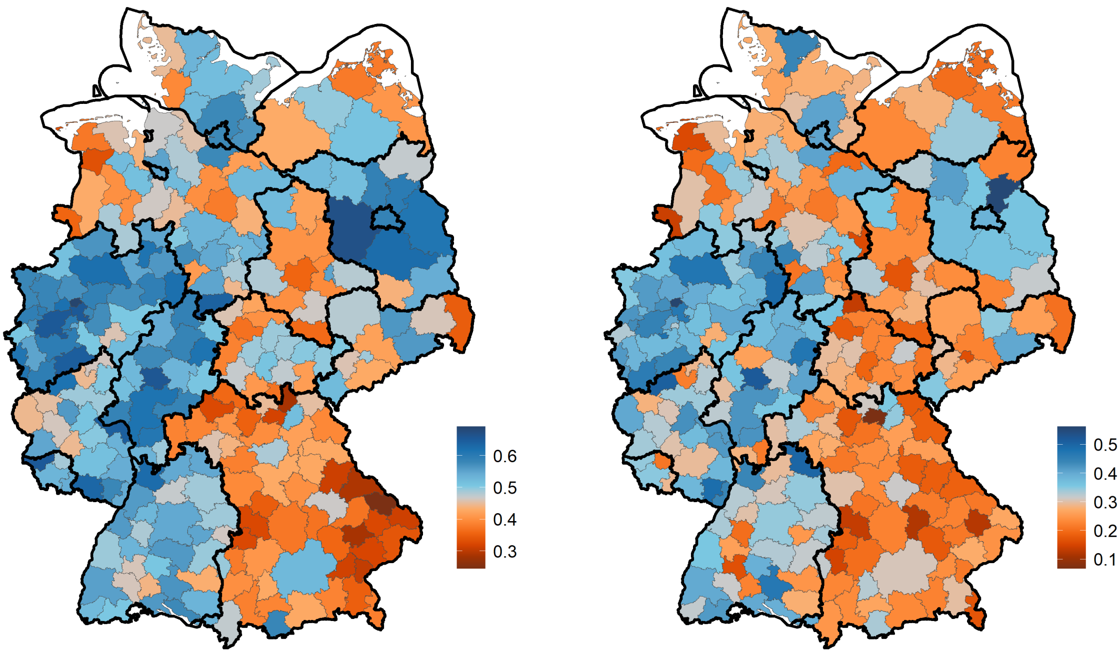
5.3 Local Labor Markets

We finally disaggregate our data once more to the level of local labor markets (LLMs). The 258 LLMs in Germany represent aggregations of counties based on commuting flows, comparable to the commuting zones in the US. Except for five local labor markets (Bremen, Bremerhaven, Hamburg, Mannheim and Ulm), all counties aggregated into LLMs belong to a single state. 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).

Regional Patterns in Absolute Mobility. We begin our local labor market-level analysis by studying regional variation in absolute mobility. Figure 9 shows the A-Level Share (Panel A) and our estimate of the Q1 measure (Panel B) in each of the 258 LLMs. Red areas correspond to regions with low, and blue areas to regions with high values of the respective statistic. For both statistics, state-level clusters are clearly visible. Panel (A) shows that the A-Level share is uniformly higher in the local labor markets of states with high average A-Level capacities, such as North Rhine-Westphalia or Hesse. Comparing the two panels demonstrates that, unsurprisingly, our measure of absolute mobility is closely linked to the local A-Level share. Consequently, we observe lower levels of absolute mobility in regions with low A-Level shares, such as Bavaria.

Overall, there exists 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 44% of the variation in the Q1 measure and 57% of the variation in the A-Level share can be attributed to state level differences.

FIGURE 9. A-Level Share and Q1 Measure by Local Labor Market



(A) A-Level Share

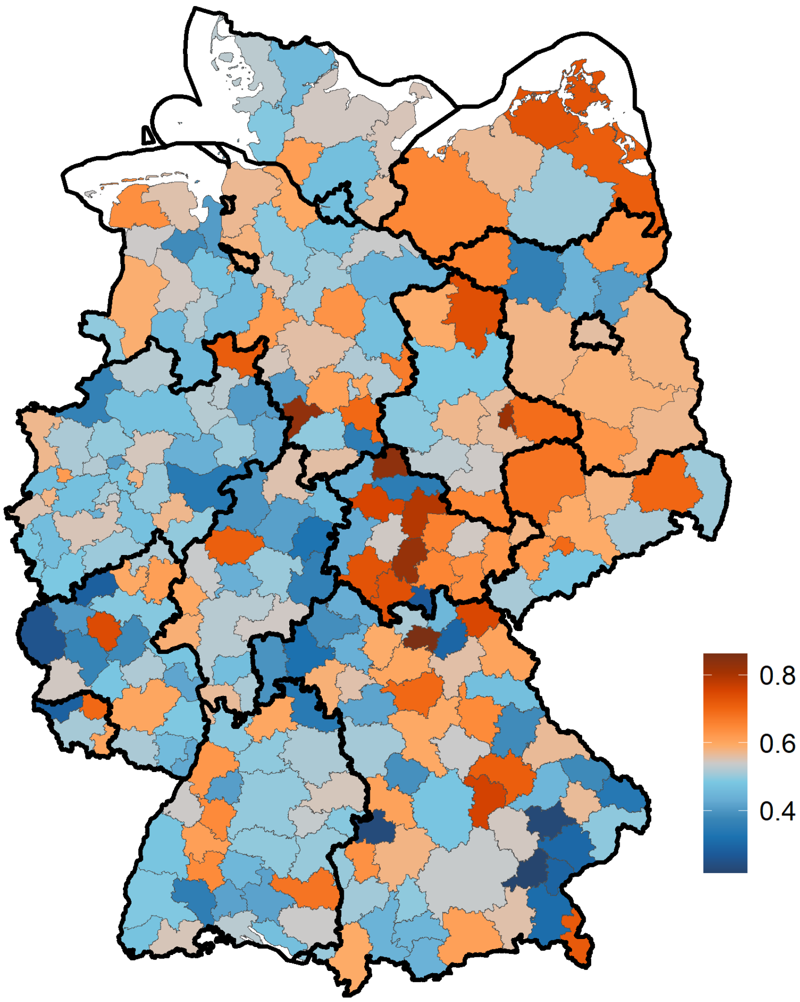
(B) Q1 Measure

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 upper stage of an A-Level track or have already attained an A-Level degree. The Q1 measure reports this same share for children in the bottom 20% of the parental income distribution.

Regional Patterns in Relative Mobility. While the variation in absolute mobility can be well explained by state A-Level shares, regional patterns in relative mobility are less obvious. Figure 10 presents a heat map of our estimates of the parental income gradient.[21] Blue areas represent regions of high mobility (low gradients), whereas red areas indicate low mobility. In some rural labor markets, the parental income gradient is estimated below 0.3, whereas in the least mobile areas the gradient exceeds 0.8. While LLMs in the East exhibit lower mobility on average, clusters of high and low mobility are spread out across all of Germany. In contrast to our estimates of absolute mobility, some of the observed clusters extend beyond state borders. The LLMs with the highest gradient (Lichtenfels) and the lowest gradient (Mühldorf) are both located in Bavaria.

[21]The corresponding heat map for the Q5/Q1 ratio is displayed in Figure B.8 in the Appendix. The correlations between our mobility measures are reported in Appendix Table B.1.

FIGURE 10. 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.

Indeed, we find that only 13% of the variation across LLMs can be explained by state level differences.

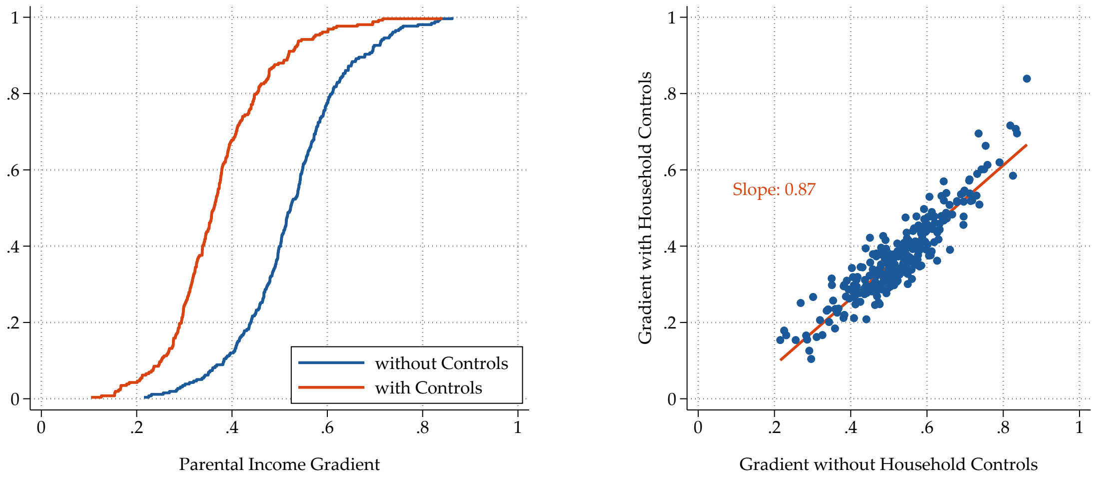
Robustness of Regional Estimates. While disaggregating our data to the LLM level allows us to ask several interesting questions, it makes it harder to distinguish meaningful variation from sampling error, as our mobility estimates are based on fewer observations. Reassuringly, the main patterns described above become also evident when computing mobility statistics at the level of spatial planning regions, essentially a higher-level aggregation of LLMs. The median number of observations per spatial planning region is 1741 (mean: 2406). Figure B.9 displays heat maps of our mobility statistics for all 96 spatial planning regions of Germany. By construction, dispersion in mobility estimates is more muted as we move to a higher level of aggregation. Yet, we still find substantial variation in mobility estimates and clusters of high and low relative mobility crossing state borders (Panel C). Moreover, it is again the case that state level differences explain more of the variation in absolute than relative mobility (72% vs 37%).

Furthermore, we show in Figure B.10 that mobility estimates for local labor markets remain virtually unchanged when computing parental income ranks not with respect to the national income distribution but with respect to the income distribution in the respective state or region type.

Sorting. What can we learn from the estimated regional differences across local labor markets? A first insight relates to the debate on the potential of place-based mobility policies. An active literature argues that places shape economic outcomes and that place-based policies can be an effective and cost-efficient tool to improve outcomes by amending local conditions (Kline and Moretti, 2014; Neumark and Simpson, 2015). In the context of educational policies and social mobility, it is often argued that the government should allocate additional resources to the local public school systems of socially immobile regions to enhance mobility. However, such a policy is unlikely to achieve its objective if social mobility in the respective regions is low for reasons other than the quality of local schools. For example, if a region exhibits a high degree of inequality in parental educational attainment, the patterns we document in Section 4.1 would likely result in low levels of relative mobility as measured by the parental income gradient.

Such systematic sorting mechanisms are at the center of the academic debate regarding the interpretation of the regional differences in estimated mobility measures within countries. For example, Rothbaum (2016) and Gallagher et al. (2018) suggest that in the US a substantial share of the geographic variation in the intergenerational mobility measures reported in Chetty et al. (2014) can be explained by differences in household characteristics across commuting zones. Unfortunately, this cannot be di-

FIGURE 11. Sorting: Conditional and Unconditional Rank Gradients



(A) Cumulative Distribution Function

(B) Scatter Plot

Notes: This figure compare unconditional and conditional estimates of the parental income gradient by local labor market. The conditioning variables include age and gender of the child, migration background, 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. Panel (A) plots the Cumulative Distribution Function (CDF) of the conditional and unconditional parental income gradients, Panel (B) shows a scatter plot of the point estimates as well as their linear fit.

rectly tested in the administrative tax data used in Chetty et al. (2014), as it contains only limited information on household characteristics.

The German census data allows us to directly test whether regional differences are muted once we account for household characteristics. We do so by computing conditional rank gradients, which we then compare to our parental income gradient. The set of conditioning variables we use for this exercise includes age and gender of the child, migration background, 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. Figure 11, Panel (A) plots the marginal distributions of conditional and unconditional rank gradients. It shows that the CDF of the unconditional gradient first order stochastically dominates the CDF of the conditional gradient, which is expected given the patterns document in Table 5. At the same time, the variance of the distribution of conditional rank gradients is approximately the same as the variance of the unconditional gradient. While this suggests that sorting does not play a major role, the same pattern would emerge if our regional estimates were dominated by sampling error, in the sense that the between local labor market variation in gradients was negligible relative to the estimation uncertainty. However, as displayed in Panel (B), we find that the relative ordering of gradients is largely unaffected by conditioning, which suggests that regional sorting of households cannot explain the regional variation in relative social mobility as we measure it. Conditional and unconditional gradients are strongly correlated, with a Pearson correlation of 0.91 and a Spearman rank correlation of 0.89. The same pattern emerges when repeating this analysis for higher levels of regional aggregation.[22]

Predictors of Mobility. If sorting cannot account for most of the spatial variation in mobility, the question remains why some regions of Germany exhibit a higher degree of social mobility than others. Similar to previous studies which document geographic variation in intergenerational mobility, we are not able to test existing theories of intergenerational transmission processes which could explain these patterns conclusively in our data. To nevertheless learn from our estimates, we conduct a prediction exercise to characterize mobile regions in more detail.

To avoid overfitting our data, we rely on a two-step approach where we preselect variables using a Random Forest before estimating a multiple linear model. Appendix C explains this procedure in more detail. We construct a comprehensive database of 73 regional indicators for this analysis, with information on labor market participation, economic conditions, infrastructure, demographics, local educational institutions and social characteristics.[23] The set of the 15 most informative predictors is displayed in Table 8, ranked by a measure of variable importance computed by the Random Forest.[24] The last column displays the sign of the bivariate correlation between each variable and the parental income gradient. A positive sign implies that the indicator predicts low mobility (a high gradient). For example, LLMs with a high prevalence of school dropouts are associated with low relative mobility. Overall, our selection procedure highlights social characteristics, the local organization of the education system and labor market conditions. These correlational findings are consistent with causal

[22] At the level of spatial planning regions, the Pearson correlation is 0.90 and the Spearman rank correlation 0.86. At the state level, the Pearson correlation amounts to 0.91 and the Spearman rank correlation to 0.84.

[23] Table C.1 lists all indicators as well as their respective sources.

[24] The exact ranking of predictors varies for different implementations of the Random Forest algorithm. We are therefore cautious not to over-interpret the ranking between single predictors.

studies that emphasize the importance of local characteristics for child and adolescent outcomes (Chetty and Hendren, 2018; Damm and Dustmann, 2014).

TABLE 8. The 15 Most Informative Predictors of Relative Mobility

Variable	Importance Measure	ρ
School Dropout Rate	0.85	+
Share Married	0.60	−
Teenage Pregnancies	0.42	+
Students	0.39	−
Median Income Vocational Qualification	0.18	−
Broadband Availability	0.17	+
Distance to Next College	0.15	−
Unemployment Rate	0.14	+
Gender Wage Gap	0.14	+
Share without Vocational Qualification	0.13	−
Gini Parental Income	0.08	−
Share Marginal Employment	0.07	−
Share Children 0-2 in Childcare	0.07	+
Share Social Assistance	0.07	+
Share on Vocational A-Level Track	0.07	−

Notes: This table lists the optimal predictive set of 15 regional indicators for the local labor market parental income gradient estimates, as chosen by a Random Forest based measure of variable importance (second column, displayed in multiples of 1000). The last column shows the sign of the Pearson correlation coefficient between each variable and the parental income gradient. A positive correlation implies that an indicator is predictive for low relative mobility (a high gradient).

In a second step, we regress the gradient on these 15 indicators selected by the algorithm. The results are reported in Table C.2. The signs of the coefficients mostly match those from the bivariate correlations in Table 8. Especially the school dropout rate, the share of married individuals, broadband availability and the share of students on vocational (rather than academic) A-Level tracks are significantly correlated with the parental income gradient in all specifications. In Appendix C.3, we discuss the connection of these results to existing theories of social mobility and provide more context to our findings. We also repeat the prediction exercise for the 129 largest and 129 smallest LLMs in Table C.3. While this analysis displays some interesting differences between rural and urban areas, the recurring themes are the same.

The key insight from this exercise is that the Random Forest algorithm is able to find meaningful variation in our data at the regional level, corresponding to existing theories of determinants of mobility. For example, as in our data, the school dropout rate is among the most significant negative correlates of relative mobility in the US data analyzed by Chetty et al. (2014). Similarly, characteristics of the vocational education system, an evergreen in the debate on social mobility in Germany, feature prominently in this list. In light of this evidence, it seems unlikely that the regional variation between LLMs is mainly driven by sampling error.

At the same time, our results do not necessarily imply that mobility differences originate from regional policy-variant parameters like the local school infrastructure, childcare availability or local employment conditions. Some of the predictors in Table 8, like the school dropout rate or the share of married individuals, could likewise point to the persistence of cultural norms or the existence of deep-rooted transmission parameters which are hard to capture with a contemporaneous set of regional indicators. For other outcomes of interest, research has shown that regional differences in Germany can reach far back into the past (e.g. Becker et al., 2020; Cantoni et al., 2019). We lack the statistical power for a detailed discrimination of these factors and exogenous variation to identify the causal determinants of mobility at the local level. We hope that future work will be able to build on our analysis and shed more light on these issues.

6 Conclusion

This paper provides novel empirical evidence on the level, evolution and geography of social mobility in Germany. Our measurement strategy allows for the use of large-scale census data and characterizes mobility using robust statistical measures of the association between the educational attainment of a child and its parents' relative position in the national income distribution. We find that on average a 10 percentile increase in parental income rank is associated with a 5.2 percentage point increase in the probability to obtain an A-Level degree, implying a top-bottom gap of approximately 50 percentage points. This gap remained stable for the 1980-1996 birth cohorts, despite a concurrent massive roll-out of higher secondary education. 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, we find that absolute mobility increased substantially.

We further document variation in mobility measures across regions and show that household characteristics cannot account for these differences. As such, our findings are consistent with place-based rather than sorting-type explanations of geographic dispersion in mobility measures. Obtaining an optimal set of mobility predictors based on our disaggregated estimates, 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 measurement approach described in this paper provides a timely and feasible way to monitor the development of social mobility in Germany for recent cohorts. This 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).

References

ACCIARI, P., A. POLO, and G. L. VIOLANTE (2019). "'And Yet, It Moves': Intergenerational Mobility in Italy". *NBER Working Paper Series* 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". *Unpublished Manuscript*.

ATTANASIO, O., T. BONEVA, and C. RAUH (2020). "Parental Beliefs about Returns to Different Types of Investments in School Children". *Journal of Human Resources* (published online ahead of print September 11, 2020).

BECKER, G. S. and N. TOMES (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility". *Journal of Political Economy* 87 (6), 1153–1189.

BECKER, S. O., L. MERGELE, and L. WOESSMANN (2020). "The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism". *Journal of Economic Perspectives* 34 (2), 143–171.

BERTRAND, M., M. MOGSTAD, and J. MOUNTJOY (2021). "Improving Educational Pathways to Social Mobility: Evidence from Norway's Reform 94". *Journal of Labor Economics* 39 (4), 965–1010.

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.

BLACK, S. and P. DEVEREUX (2011). "Recent Developments in Intergenerational Mobility". *Handbook of Labor Economics*. Ed. by O. ASHENFELTER and D. CARD. Vol. 4. Elsevier. Chap. 16, 1487–1541.

BLACK, S. E., P. J. DEVEREUX, P. LUNDBORG, and K. MAJLESI (2020). "Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviors". *Review of Economic Studies* 87 (4), 1683–1725.

BOAR, C. and D. LASHKARI (2021). "Occupational Choice and the Intergenerational Mobility of Welfare". *NBER Working Paper Series* 29381.

BONEVA, T. and C. RAUH (2021). "Socio-economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns". *Unpublished Manuscript*.

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.

BÜTIKOFER, A. and G. PERI (2021). "How Cognitive Ability and Personality Traits Affect Geographic Mobility". *Journal of Labor Economics* 39 (2), 559–595.

CANTONI, D., F. HAGEMEISTER, and M. WESTCOTT (2019). "Persistence and Activation of Right-Wing Political Ideology". *CRC Discussion Paper* 143.

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

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 Discussion Paper* 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.* Hamburg: Nannen-Verlag.

DAMM, A. P. and C. DUSTMANN (2014). "Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?" *American Economic Review* 104 (6), 1806–1832.

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

DUSTMANN, C. (2004). "Parental Background, Secondary School Track Choice, and Wages". *Oxford Economic Papers* 56 (2), 209–230.

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.

FUCHS-SCHÜNDELN, N., D. KRUEGER, and M. SOMMER (2010). "Inequality Trends for Germany in the Last Two Decades: A Tale of Two Countries". *Review of Economic Dynamics* 13 (1), 103–132.

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". *Die Bildungsexpansion*. VS Verlag für Sozialwissenschaften., 11–24.

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.

HELBIG, M. and R. NIKOLAI (2015). *Die Unvergleichbaren: Der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949*. Verlag Julius Klinkhardt.

HILGER, N. G. (2015). "The Great Escape: Intergenerational Mobility in the United States Since 1940". *NBER Working Paper Series* 21217.

KALLEBERG, A. L. (1977). "Work Values and Job Rewards: A Theory of Job Satisfaction". *American Sociological Review*, 124–143.

— (2011). *Good Jobs, Bad Jobs: The Rise of Polarized and Precarious Employment Systems in the United States, 1970s-2000s*. Russell Sage Foundation.

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.

LANDERSØ, R. and J. J. HECKMAN (2017). "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US". *Scandinavian Journal of Economics* 119 (1), 178–230.

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

LIANG, K.-Y. and S. L. ZEGER (1986). "Longitudinal Data Analysis Using Generalized Linear Models". *Biometrika* 73 (1), 13–22.

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.

— (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.

MOTTAZ, C. J. (1985). "The Relative Importance of Intrinsic and Extrinsic Rewards as Determinants of Work Satisfaction". *Sociological Quarterly* 26 (3), 365–385.

MUÑOZ, E. (2021). "The Geography of Intergenerational Mobility in Latin America and the Caribbean". *Stone Center On Socio-Economic Inequality Working Paper Series* 29.

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

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.

RAUH, C. (2017). "Voting, Education, and the Great Gatsby Curve". *Journal of Public Economics* 146, 1–14.

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*.

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.

STATISTISCHES BUNDESAMT (2018). *Mikrozensus 2018. Qualitätsbericht*. Wiesbaden.

— (2021). *Konsumausgaben von Familien für Kinder: Berechnungen auf der Grundlage der Einkommens- und Verbrauchsstichprobe 2018*. Wiesbaden, 31.

A Additional Information on the Mikrozensus

The Microcensus (Mikrozensus, MZ) is the largest household survey in Europe. Conducted annually with a sampling fraction of 1% of all individuals who have the right of residence in Germany, it yields representative statistics on the German population. The MZ has been conducted in West Germany since 1957 and in the new federal states (East Germany) since 1991. It is planned and prepared by the Federal Statistical Office of Germany and carried out by the statistical offices of the 16 German states. The legal basis of the MZ is the Microcensus Law, which makes it compulsory for households to provide answers to the core items of the survey. The non-response rate is further minimized by repeated visits of interviewers to non-responding households and multiple possible ways for the sampled households to submit information.

FIGURE A.1. Illustration of the Microcensus Survey Design

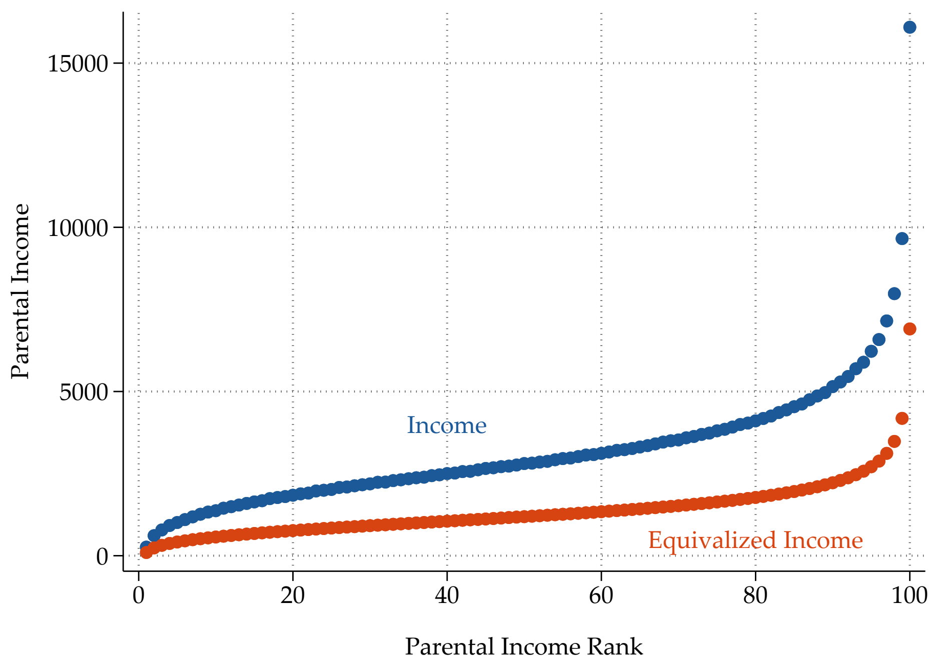
	Rotation Quarter							
Survey Wave	1	2	3	4	5	6	7	···
1	✓	✓	✓	✓	X	X	X	···
2	X	✓	✓	✓	✓	X	X	⋱
3	X	X	✓	✓	✓	✓	X	⋱
4	X	X	X	✓	✓	✓	✓	⋱
⋮	⋮	⋱	⋱	⋱	⋱	⋱	⋱	⋱

Since 1972, the MZ uses a single-stage stratified cluster sampling design. The primary sampling units typically consist of neighboring buildings (larger buildings are divided into smaller partitions). For the survey waves utilized in this paper, the target size for a cluster is 7-15 households. All households and residents in the sampled clusters are interviewed. The database used to assign households to clusters is created based on the most recent full census and updated annually using information on new construction activities. Since 1977, each cluster is assigned to a "rotation quarter" that remains in the survey for four years. Each year, a quarter is replaced by new clusters. The survey does not follow individuals who leave their cluster, but replaces them by the new residents. The MZ survey design results in data best described as a repeated survey with partial overlap of units, as sketched in Figure A.1.

Due to data protection laws, we do only observe this panel structure in our data following wave 2011. In Section 4.2, we therefore cannot cluster standard errors at the level of time-constant primary sampling units. We instead cluster standard errors at the household level. As the number of households per cluster is low, the consequences for standard errors are negligible.

Sample Income Distribution and Ranks. Figure A.2 displays the sample distribution of equivalized monthly net household income and the corresponding percentile ranks in the 2011-2018 MZ data. We CPI adjust all household incomes in order to allow for meaningful aggregation of survey-years before computing ranks. Ties are broken by allocating households to the lower quantile. Our findings are insensitive to the choice of tie-breakers. Ranks are computed separately for each year within the sample of all households that have at least one co-resident child in the age range 17-21.

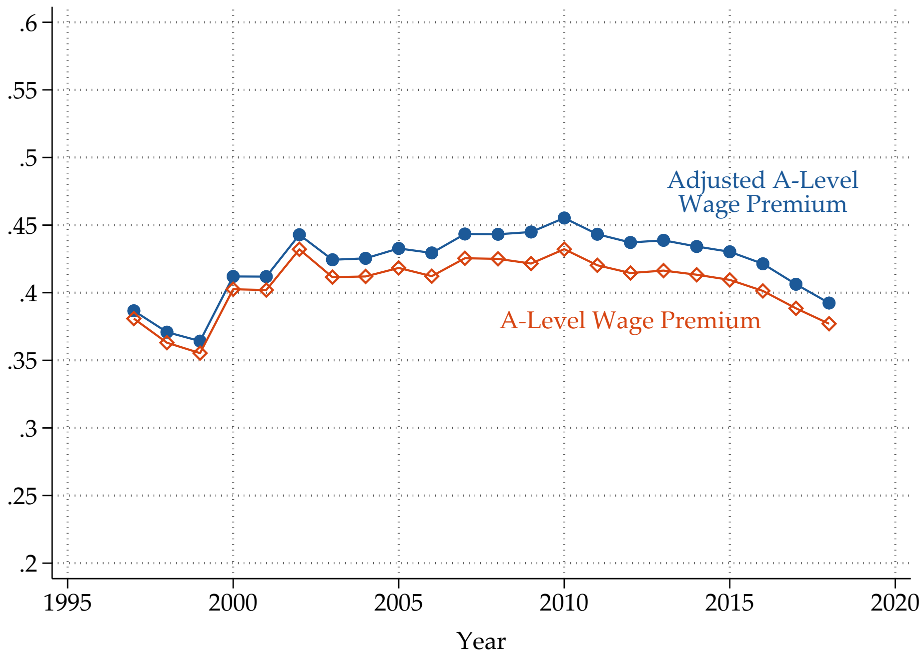
FIGURE A.2. Household Income by Percentile Rank



Notes: This figure plots equivalized net monthly household income (net of income of dependent children) by parental income rank in the 2011-2018 MZ data. Equivalization is based on the modified OECD scale. For comparison, the non-equivalized values are plotted as well. Both income measures are expressed in constant 2015 Euro.

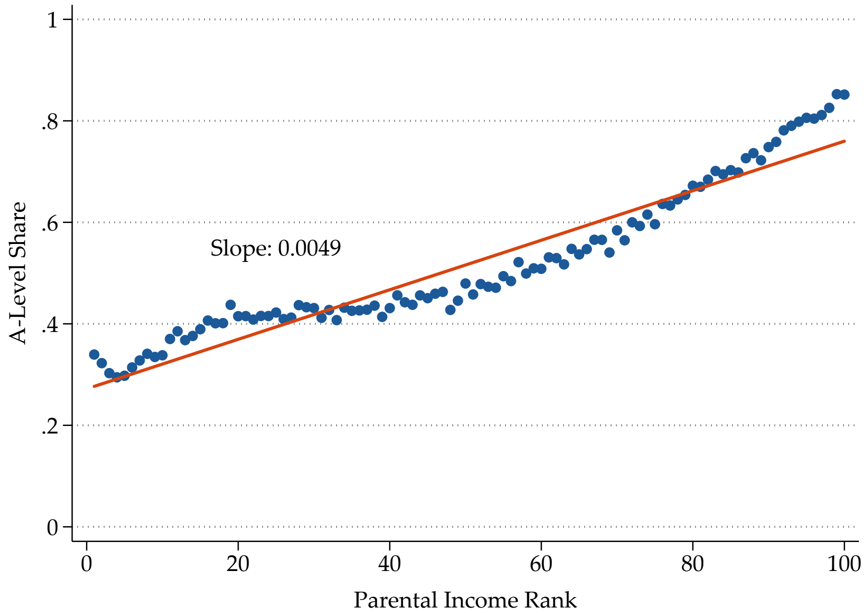
B Additional Figures and Tables

Figure B.1. A-Level Wage Premium, Years 1997-2016

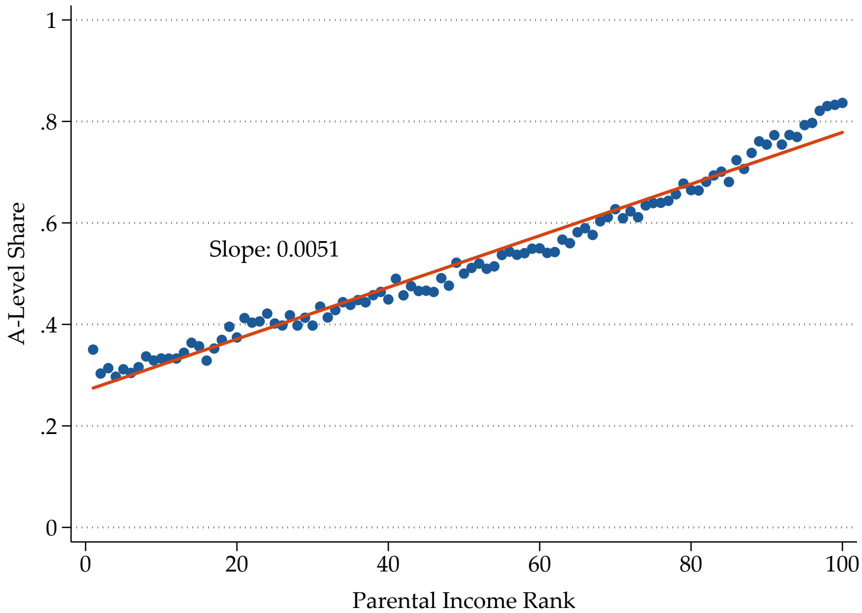


Notes: This figure shows the development of the A-Level wage premium for the years 1997-2016 as computed in the MZ. We compute the A-Level wage premium by regressing the log of net monthly personal income of full-time working employees aged 30-45 on an A-Level dummy. The adjusted A-Level wage premium is computed by additionally conditioning on a set of age indicators to indirectly account for job experience.

FIGURE B.2. National Estimates under Different Equivalization Schemes



(A) No Adjustment

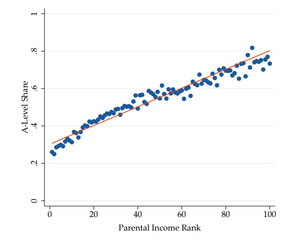


(B) Per Capita Adjustment

Notes: This figure shows the fraction of children aged 17-21 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree by percentile rank of their parents in the national income distribution based on the MZ waves 2011-2018, as well as the best linear approximation to the empirical CEF. In Panel (A), parental income is not adjusted for household size, whereas in Panel (B) income is divided by the number of household members. The OLS slopes reported in the figure are estimated using the underlying micro data.

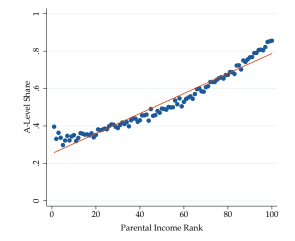
FIGURE B.3. Social Mobility for Subgroups

(A) Single Parent

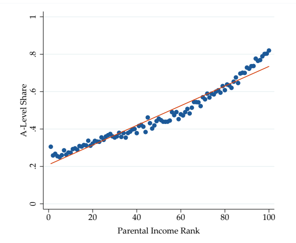


(B) Two Parents

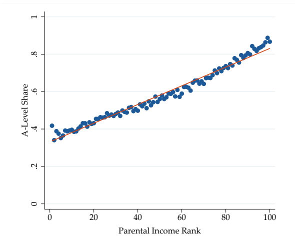
(C) Parents Not Married



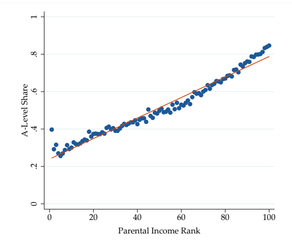
(D) Parents Married



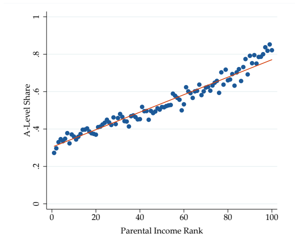
(E) Male



(F) Female

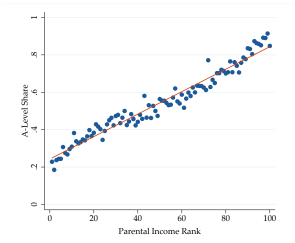


(G) Native

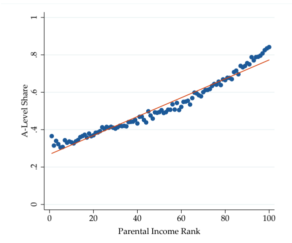


(H) Migrant

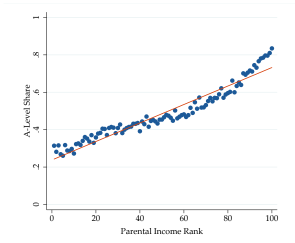
(I) West Germany



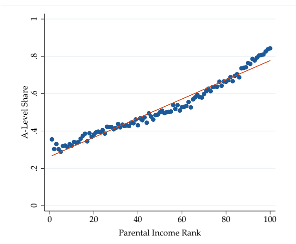
(J) East Germany



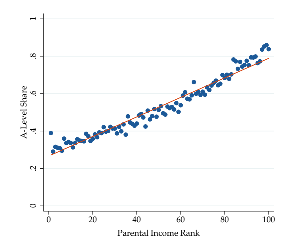
(K) Siblings



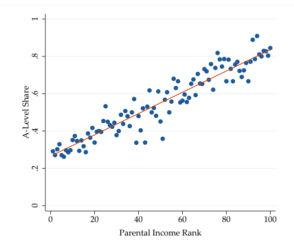
(L) No Siblings



(M) First Child



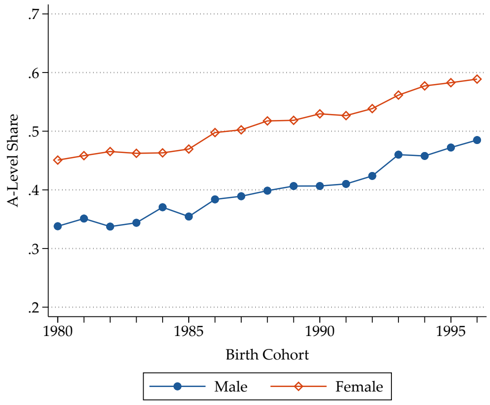
(N) Second Child



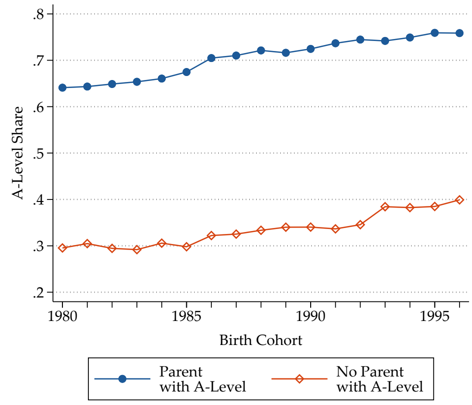
(O) Third Child or Later

Notes: This figure shows for different population subgroups the fraction of children aged 17-21 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree by percentile rank of their parents in the national income distribution based on the MZ waves 2011-2018, as well as the best linear approximation to the empirical CEF. 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.

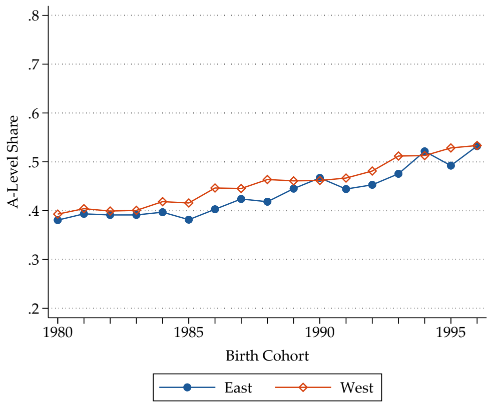
FIGURE B.4. Time Trend A-Level Share for Subgroups



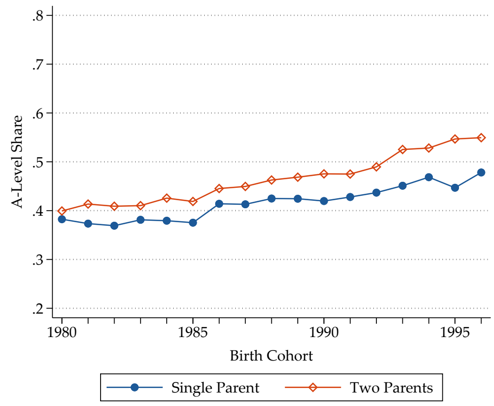
(A) Gender



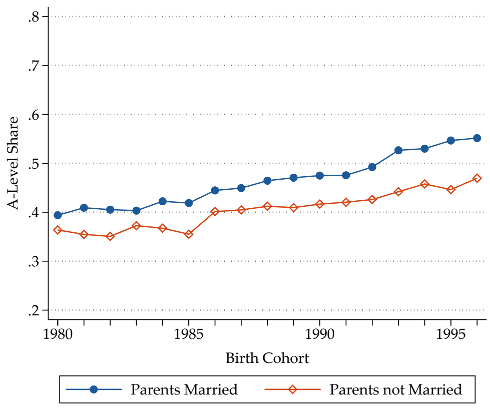
(B) Parental Education



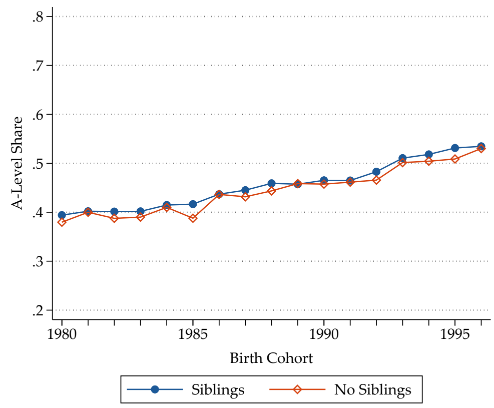
(C) Region



(D) Parenting Status



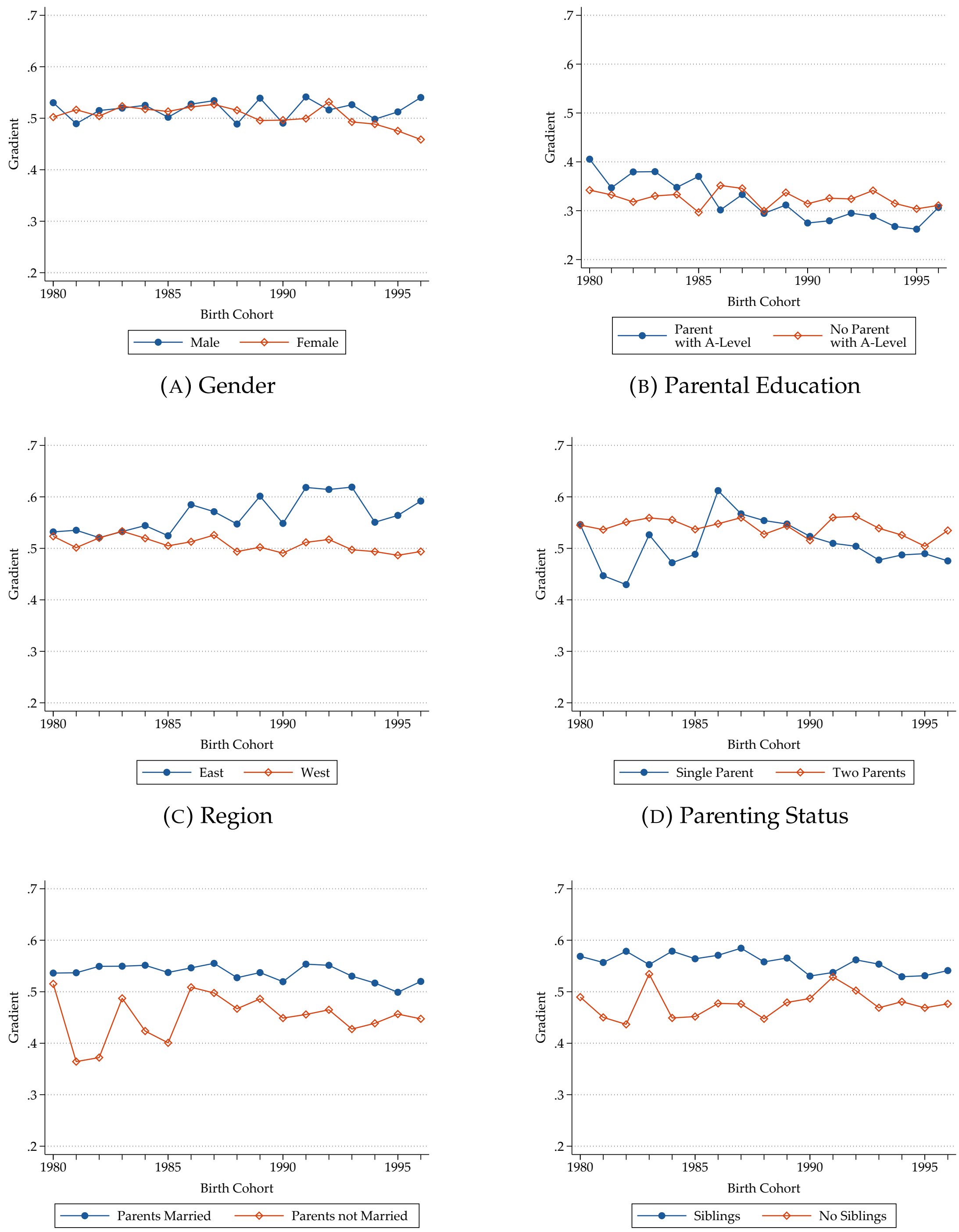
(E) Parental Marital Status



(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 upper stage of an A-Level track or have already attained an A-Level degree.

FIGURE B.5. Time Trend Parental Income Gradient for Subgroups



(A) Gender

(B) Parental Education

(C) Region

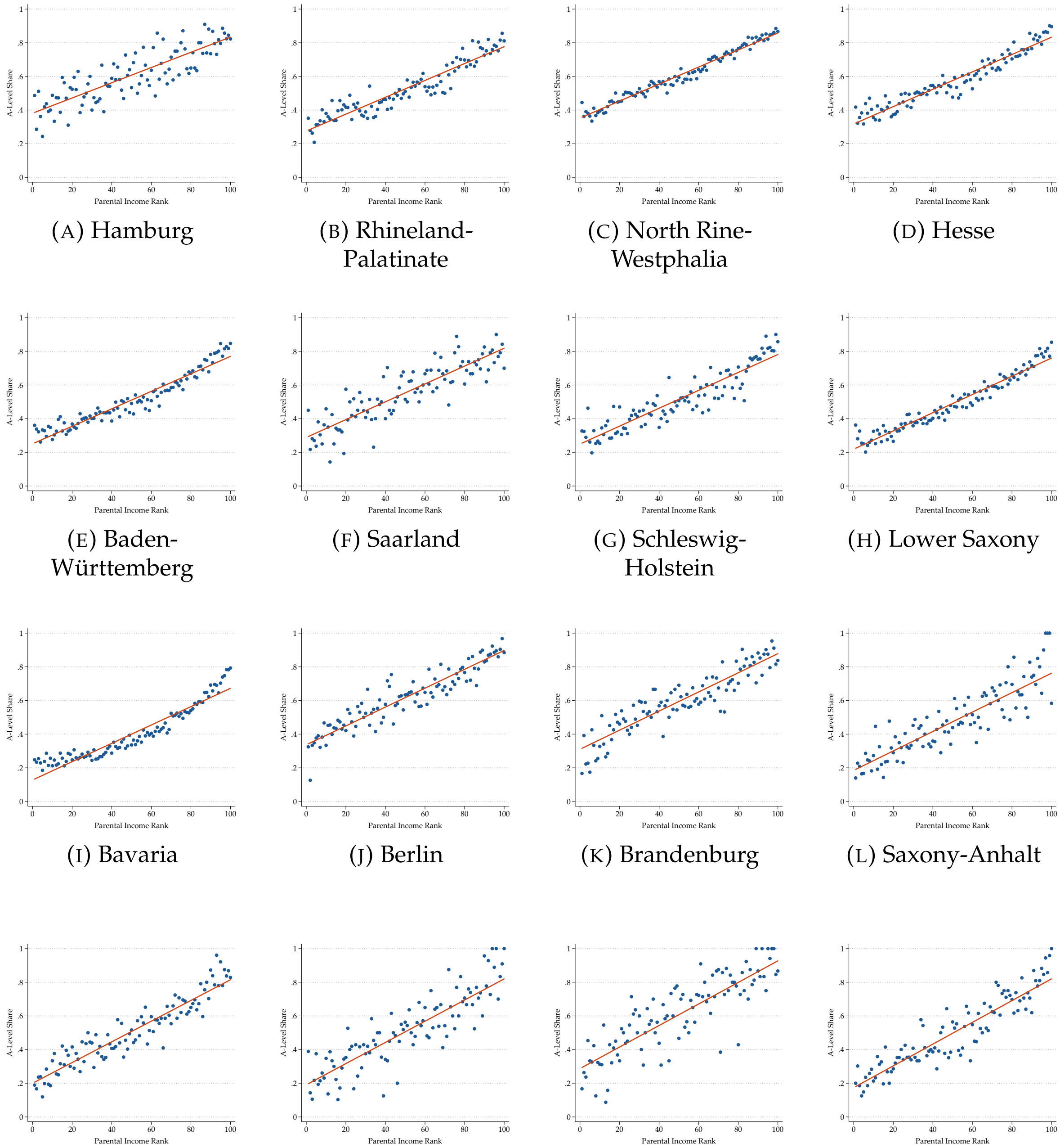
(D) Parenting Status

(E) Parental Marital Status

(F) Siblings

Notes: This figure shows the development of the parental income gradient for different population subgroups for birth cohorts 1980-1996 in the MZ. The parental income gradient per cohort is estimated as $100 \times \gamma_t$ in 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.

FIGURE B.6. Social Mobility at the State Level



(A) Hamburg

(B) Rhineland-Palatinate

(C) North Rine-Westphalia

(D) Hesse

(E) Baden-Württemberg

(F) Saarland

(G) Schleswig-Holstein

(H) Lower Saxony

(I) Bavaria

(J) Berlin

(K) Brandenburg

(L) Saxony-Anhalt

(M) Saxony

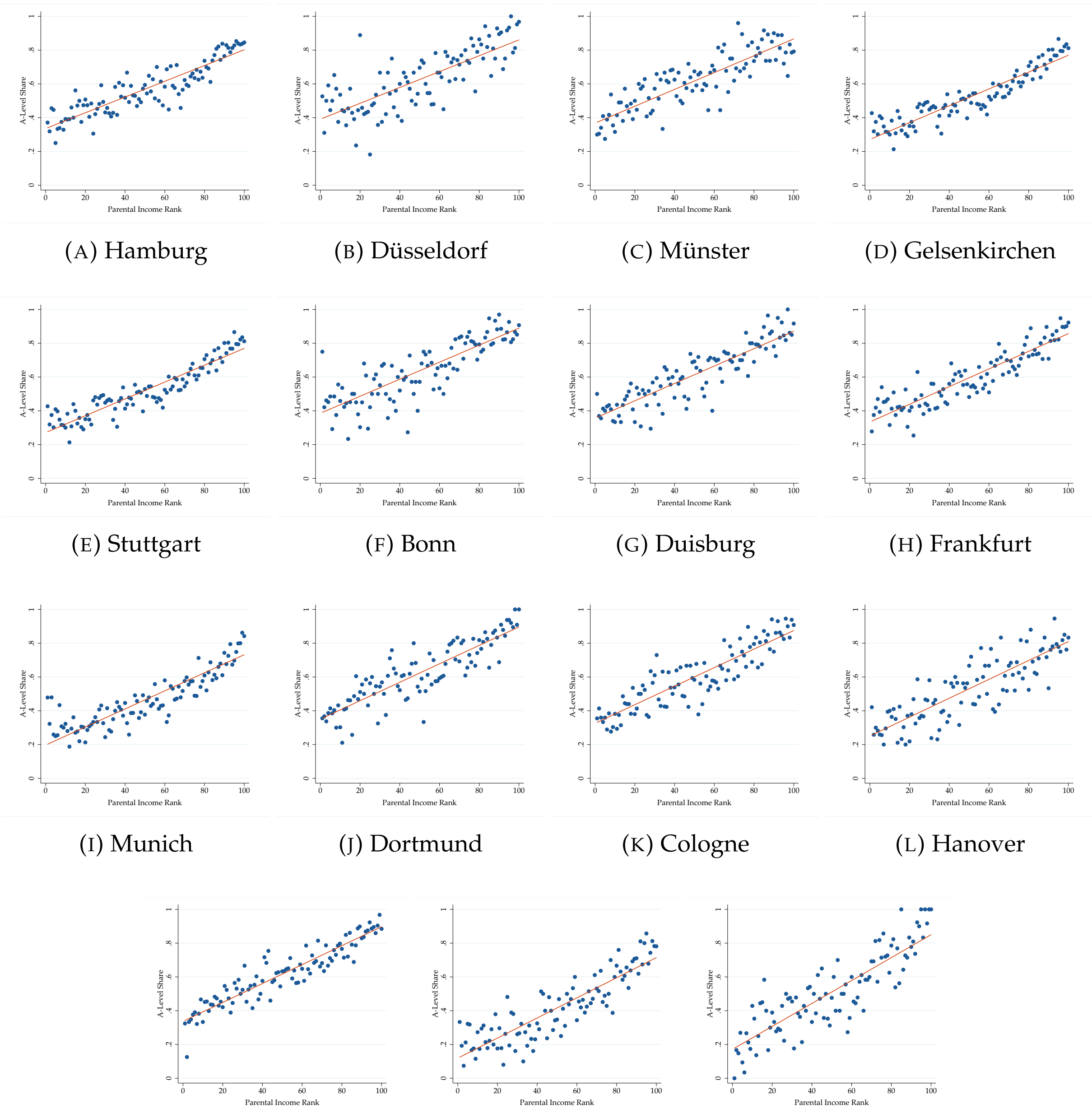
(N) Mecklenburg-Vorpommern

(O) Bremen

(P) Thuringia

Notes: This figure shows for each German state the fraction of children aged 17-21 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree by percentile rank of their parents in the national income distribution based on the MZ waves 2011-2018, as well as the best linear approximation to the empirical CEF.

FIGURE B.7. Social Mobility for Cities



(A) Hamburg (B) Düsseldorf (C) Münster (D) Gelsenkirchen

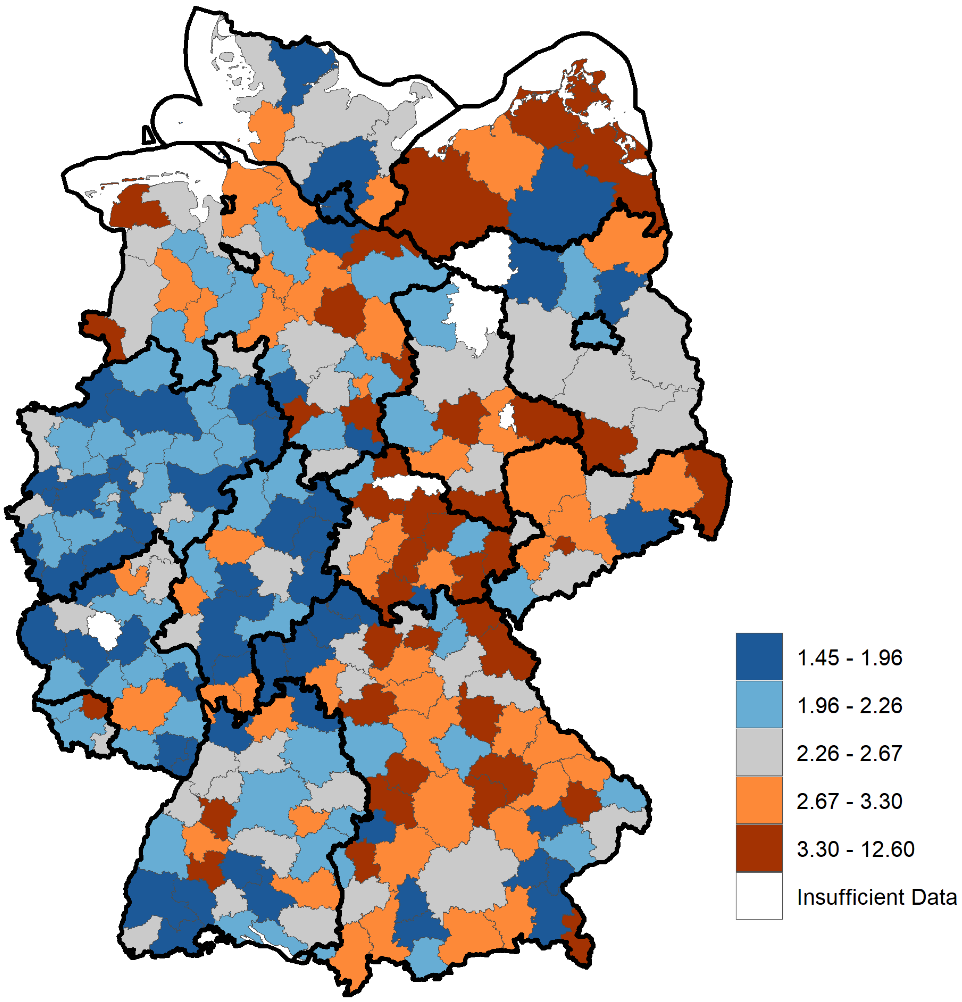
(E) Stuttgart (F) Bonn (G) Duisburg (H) Frankfurt

(I) Munich (J) Dortmund (K) Cologne (L) Hanover

(M) Berlin (N) Nuremberg (O) Leipzig

Notes: This figure shows for the 15 largest (by population size in 2017) local labor markets in Germany the fraction of children aged 17-21 that are either enrolled in the upper stage of an A-Level track or have already attained an A-Level degree by percentile rank of their parents in the national income distribution based on the MZ waves 2011-2018, as well as the best linear approximation to the empirical CEF.

FIGURE B.8. **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. The colors indicate the quintile of the respective LLM point estimate in the distribution of estimates to account for outliers of the Q5/Q1 ratio induced by small denominators. 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.

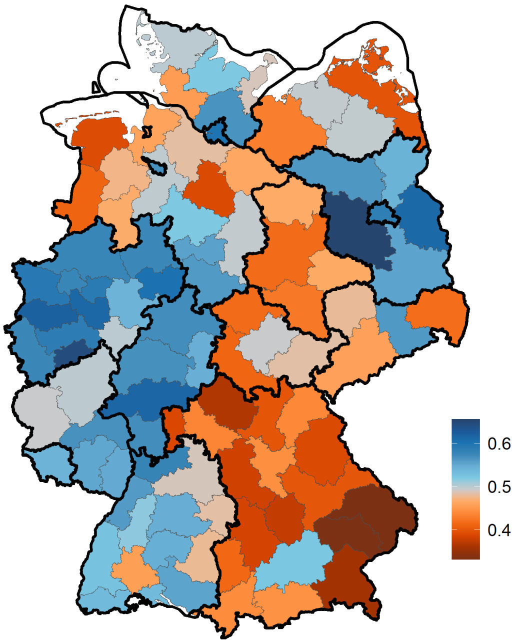
TABLE B.1. Correlation between Mobility Measures

Measure	Corr.	A-Level	Q1	Q5/Q1	Gradient
A-Level	ρ	1	-	-	-
	r	1	-	-	-
Q1	ρ	0.76	1	-	-
	r	0.78	1	-	-
Q5/Q1	ρ	-0.40	-0.72	1	-
	r	-0.48	-0.84	1	-
Gradient	ρ	-0.01	-0.45	0.65	1
	r	-0.07	-0.47	0.76	1

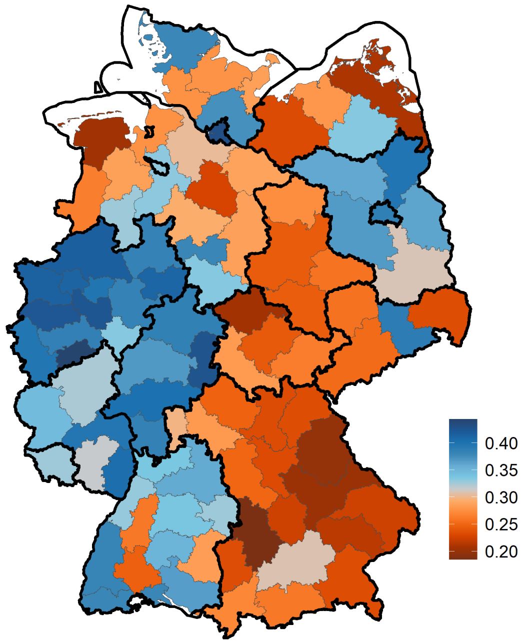
Notes: This table reports the pairwise correlations between estimates of different measures of social mobility across LLMs in Germany. ρ denotes the Pearson correlation coefficient, r denotes the Spearman rank correlation coefficient.

Table B.1 reports the correlations between our mobility measures. While the Q1 measure is well predicted by the unconditional A-Level share, there exists no systematic association between the A-Level share and the parental income gradient, highlighting that the gradient is not sensitive to the baseline probability of obtaining an A-Level degree. Finally, the correlation between the parental income gradient and the Q1 measure ranges below -0.5, demonstrating that a high level of absolute mobility in a given LLM does not always imply a high level of relative mobility.

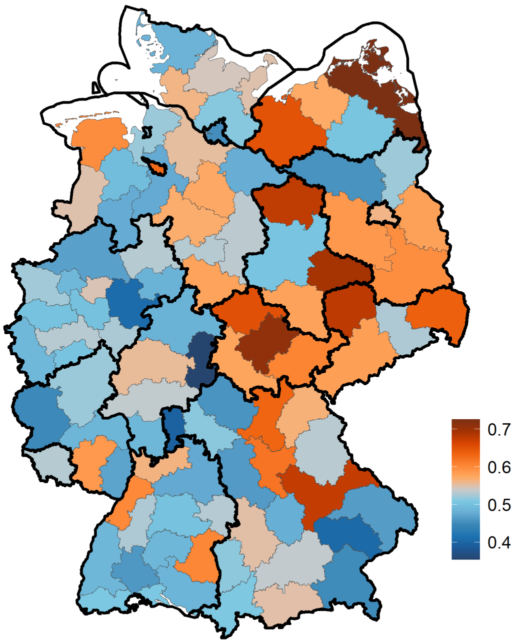
FIGURE B.9. Mobility Estimates by Spatial Planning Region



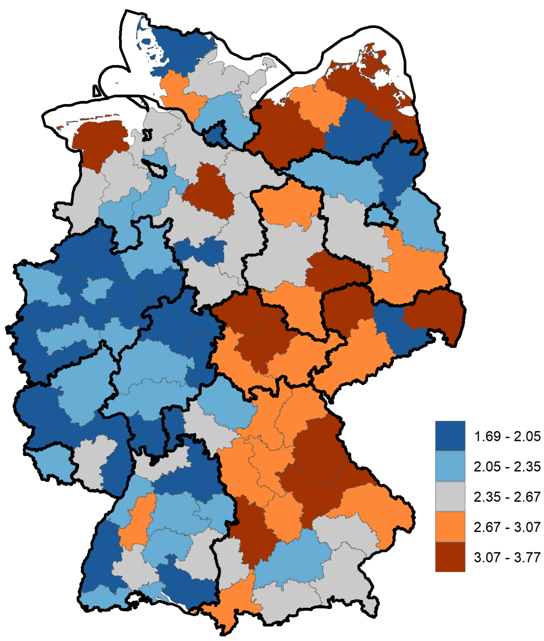
(A) A-Level Share



(B) Q1 Measure



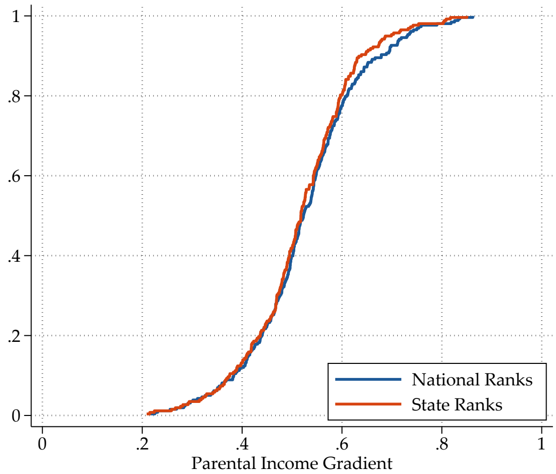
(C) Gradient



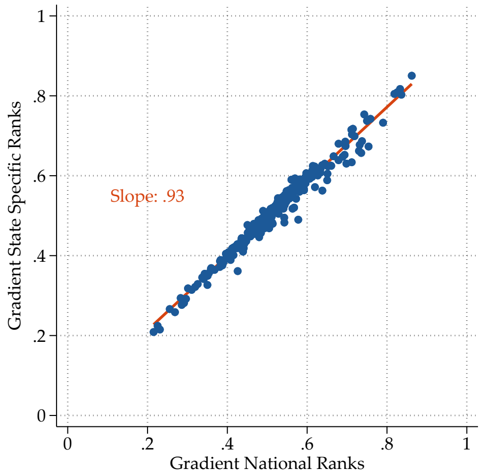
(D) Q5/Q1 Ratio

Notes: This figure presents heat maps of the A-Level share (Panel A), the Q1 measure (Panel B), the parental income gradient (Panel C) and the Q5/Q1 ratio (Panel D) for the 96 spatial planning regions of Germany. Spatial planning regions constitute a more comprehensive version of the LLMs discussed in Section 5, as they also represent aggregations of counties based on commuting flows. Children are assigned to spatial planning regions 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. In Panel (D), the colors indicate the quintile of the respective point estimate in the distribution of estimates to account for outliers of the Q5/Q1 ratio induced by small denominators.

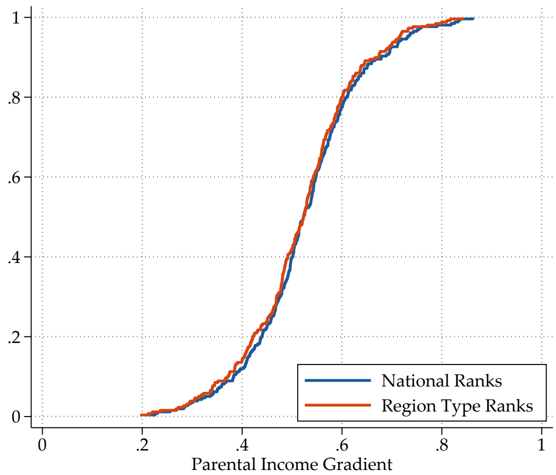
FIGURE B.10. Robustness to State and Region Specific Parental Income Ranks



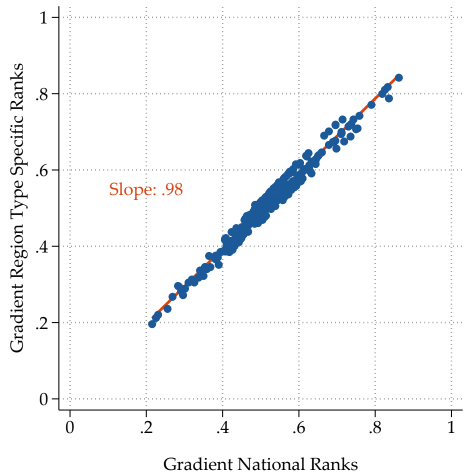
(A) CDFs



(B) Scatter Plot National/State



(C) CDFs



(D) Scatter Plot National/Region Type

Notes: This figure displays the sensitivity of our LLM-level estimates of the parental income gradient with respect to the reference income distribution. For this aim, the upper two panels compare gradients computed based on the national and the state-specific income distributions: Panel (A) displays the Cumulative Distribution Function (CDF) of both gradients, Panel (B) shows a scatter plot of the point estimates as well as their linear fit. The bottom two panels compare the gradients obtained by computing income ranks based on the national and the region-type-specific income distribution. The region types are defined by the Federal Institute for Building, Urban Affairs and Spatial Research (BBSR) and classify each county into either urban, suburban or rural. For LLMs comprising of counties of different types, we assign the most frequent category. Again, Panel (C) displays the Cumulative Distribution Function (CDF) of both gradients, whereas Panel (D) shows a scatter plot of the point estimates as well as their linear fit. The reported slope parameters of 0.93 and 0.98 correspond to the OLS slope estimates obtained by regressing the gradients computed by using the respective local ranks on the gradients computed by using national income ranks.

C Regional Predictors of Mobility

C.1 Regional Indicators

Table C.1 displays all 73 regional indicators we use as predictors in the Random Forest algorithm. In a first step, we retrieve data from the 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 we suppose to be potentially relevant for social mobility and are not collinear: 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, the share of A-Level degrees obtained on vocational schools 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 regional statistics on the LLM level using 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 variables by averaging the local indicators over the years 2011-2018 at the LLM level.

TABLE C.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 Students Enrolled in Gymnasium	INKAR
	Share 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 A-Level Degree from Vocational Schools	Destatis
	Mean Parental Education	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
	Expected Rank Difference Parental 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 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, the Statistical Office of Germany (Destatis), the Hochschulrektorenkonferenz (HRK) or the Mikrozensus (MZ).

C.2 Prediction Exercise

To study the association between local characteristics and intergenerational mobility, prior literature has typically relied on correlation coefficients or estimated multiple linear models (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), resulting in diminished external validity. 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.[25] 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 a Lasso regression. This is because we find that a linear Lasso model fits our data poorly: 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 39% of the variation in the test sample (R^2 = 0.39), whereas the predictive power of Lasso (R^2 = 0.15) and Elastic Net (R^2 = 0.17) is lower. The results for Lasso and Elastic Net are based on λ chosen by 5-fold cross-validation. For the Random Forest, we fit 1000 trees and randomly select $73/3 \approx 24$ variables for each split.

Before constructing the Random Forest, we standardize all 73 indicators to have mean 0 and standard deviation 1. Once the Random Forest is fitted, we can rank covariates according to their predictive power and thereby obtain a 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.

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

C.3 Additional Results

Regression Estimates. The second step of our prediction exercise consists of estimating a multiple linear model, where we regress the parental income gradient on the 15 most informative indicators as chosen by the Random Forest. All right-hand side variables are standardized so that the coefficients report the association between a one standard deviation change in the covariate and an absolute change in the gradient. The results are reported in Table C.2. The signs of the coefficients often match those from the bivariate correlations in Table 8. For example, a one standard deviation increase in the school dropout rate is associated with a 3.9 percentage point higher parental income gradient. This association becomes stronger when adding state indicators. A high gradient also aligns with a high number of teenage pregnancies, a high unemployment rate and a large share of households with access to broadband Internet. 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 aged 0-2 in childcare and for the share of children on a vocational A-Level track. Due to the limited sample size of 258 local labor markets, we lack the power to precisely estimate most coefficients. Exceptions are the school dropout rate, broadband availability, the share of married individuals and the share of children on a vocational A-Level track.

TABLE C.2. Social Mobility and Regional Characteristics

	(1)	(2)	(3)	(4)	(5)
School Dropout Rate	0.0391	0.0371	0.0393	0.0554	0.0539
	(0.0110)	(0.0110)	(0.0091)	(0.0162)	(0.0162)
Share Married	-0.0225	-0.0286	-0.0225	-0.0243	-0.0278
	(0.0089)	(0.0089)	(0.0065)	(0.0108)	(0.0109)
Teenage Pregnancies	0.0169	0.0123	0.0211	0.0160	0.0115
	(0.0226)	(0.0231)	(0.0155)	(0.0252)	(0.0266)
Students	-0.0143	-0.0166	-0.0055	-0.0214	-0.0246
	(0.0131)	(0.0128)	(0.0093)	(0.0164)	(0.0165)
Median Income Vocational Qualification	-0.0179	-0.0194	-0.0025	-0.0234	-0.0224
	(0.0129)	(0.0129)	(0.0114)	(0.0167)	(0.0177)
Broadband Availability	0.0260	0.0274	0.0194	0.0231	0.0261
	(0.0100)	(0.0100)	(0.0085)	(0.0105)	(0.0109)
Distance to Next College	-0.0048	-0.0059	-0.0051	-0.0025	-0.0045
	(0.0072)	(0.0077)	(0.0070)	(0.0072)	(0.0076)
Unemployment Rate	0.0368	0.0295	0.0124	0.0537	0.0476
	(0.0365)	(0.0365)	(0.0236)	(0.0464)	(0.0470)
Gender Wage Gap	-0.0029	-0.0041	0.0048	0.0156	0.0124
	(0.0142)	(0.0144)	(0.0126)	(0.0174)	(0.0177)
Share without Vocational Qualification	0.0057	0.0085	-0.0035	0.0132	0.0108
	(0.0171)	(0.0173)	(0.0127)	(0.0217)	(0.0220)
Gini Parental Income	-0.0171	-0.0108	-0.0236	0.0051	0.0117
	(0.0147)	(0.0150)	(0.0110)	(0.0200)	(0.0209)
Share Marginal Employment	-0.0086	-0.0162	-0.0183	-0.0222	-0.0250
	(0.0138)	(0.0142)	(0.0121)	(0.0152)	(0.0154)
Share Children 0-2 in Childcare	-0.0398	-0.0420	-0.0526	-0.0246	-0.0259
	(0.0192)	(0.0189)	(0.0182)	(0.0234)	(0.0236)
Share Social Assistance	-0.0607	-0.0429	-0.0406	-0.0969	-0.0782
	(0.0343)	(0.0361)	(0.0231)	(0.0450)	(0.0498)
Share on Vocational A-Level Track	-0.0165	-0.0171	-0.0133	-0.0213	-0.0224
	(0.0092)	(0.0092)	(0.0079)	(0.0100)	(0.0100)
Additional Controls	-	✓	✓	-	✓
State Indicators	-	-	-	✓	✓
Weighted	-	-	✓	-	-
N	258	258	258	252	252
R^2	0.256	0.273	0.253	0.296	0.305

Notes: Each column of this table reports coefficients from a linear regression with robust standard 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 8) are standardized to have mean 0 and standard deviation 1. Columns (3) and (4) contain state dummies, 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 whether 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.

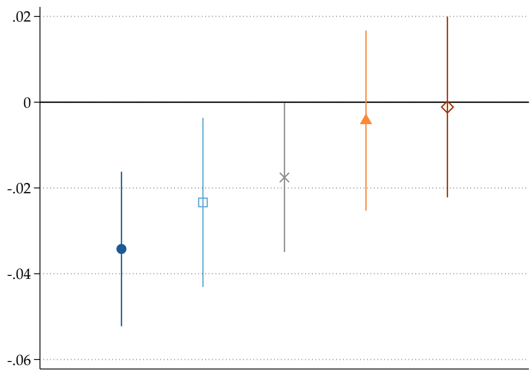
Graphical Evidence. To understand the relationship between relative mobility and the indicators with the largest t-statistics in more detail, we separately regress the A-Level share in each quintile of the parental income distribution on each indicator and plot the estimates in Figure C.1. These plots 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 C.1, 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 low socio-economic background, further information is needed to test this hypothesis in detail.[26]

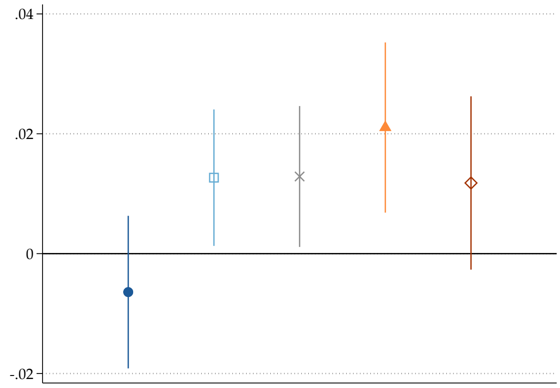
Panel (B) sheds light on the negative connection between broadband availability and mobility. While broadband access is associated with a higher A-Level share on average, this is not true for children in Q1, for whom the relationship becomes negative. We can only speculate about the reasons. Broadband access is highly correlated with factors pointing to dynamic and prosperous labor markets, which exhibit above

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

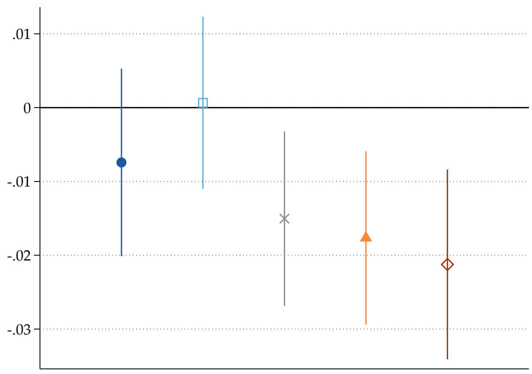
FIGURE C.1. Predicting the A-Level Share by Parental Income Quintile



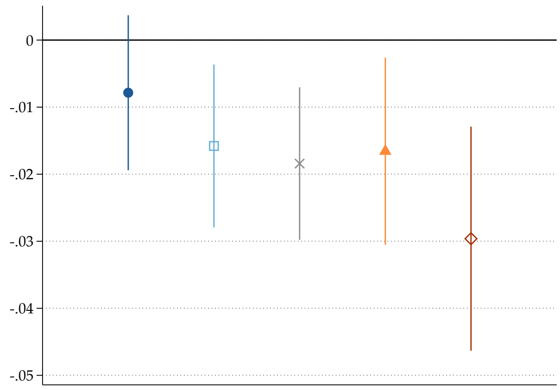
(A) School Quality



(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 indicators 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 indicators, 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.

average inequality. For that reason, broadband availability may proxy urban areas in which all but children from the bottom of the income distribution profit from a dynamic and rewarding economic environment. 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 quintile 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 on 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 effects 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 local labor markets 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 of attending the general high track (Gymnasium) at the age of 13-14 than of obtaining an A-Level degree later on (gradient of 0.55 versus 0.52), again suggesting that vocational schools may mediate the influence of parental background.

TABLE C.3. The 15 Most Informative Predictors by LLM Size

Variable	Importance Measure	ρ
Panel (A): 129 Largest Local Labor Markets		
School Dropout Rate	0.42	+
Gini Parental Income	0.23	−
Share Married	0.16	−
Share without Vocational Qualification	0.10	−
Students	0.09	−
Physician Density	0.09	+
Teenage Pregnancies	0.06	+
Mean Parental Income	0.06	−
Share Marginal Employment	0.06	−
Students (Universities of Applied Sciences)	0.05	+
Median Income Vocational Qualification	0.05	−
Distance to next Elementary School	0.03	−
Share Children 3-5 in Childcare	0.03	+
Child Mortality	0.03	−
Ratio p50/p15	0.03	−
Panel (B): 129 Smallest Local Labor Markets		
School Dropout Rate	0.75	+
Unemployment Rate	0.45	+
Child Poverty	0.40	+
Students	0.40	−
Share Married	0.33	−
Teenage Pregnancies	0.33	+
Median Income Vocational Qualification	0.19	−
Gender Wage Gap	0.19	+
Share Social Assistance	0.18	+
Total Net Migration	0.12	−
Highly Qualified Persons	0.10	+
Broadband Availability	0.10	+
Share on Vocational A-Level Track	0.08	−
Building Permits	0.08	−
Share Apartment Buildings	0.07	+

Notes: This table lists the 15 most predictive indicators for explaining variation in the parental income gradient between local labor markets in Germany, separately for the 129 largest (Panel [A]) and the 129 smallest (Panel [B]) local labor markets. See Section C.2 for details on the implementation via a Random Forest variable importance measure. The second column displays the measure of variable importance (in multiples of 1000). The last column shows the sign of the Pearson correlation coefficient between each variable and the parental income gradient. A positive correlation implies that an indicator is predictive for low relative mobility (a high gradient).

References

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.

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.

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

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.

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

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.

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.

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* 8846.

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