AERA Open
January-December 2025, Vol. 11, No. 1, pp. 1–26
DOI: 10.1177/23328584251331453
Article reuse guidelines: sagepub.com/journals-permissions

Metropolitan, Urban, and Rural Regions: How Regional Differences Affect Elementary School Students in Germany

Jakob Schwerter
TU Dortmund University

Johannes Bleher
University of Hohenheim

Philipp Doebler
Nele McElvany
TU Dortmund University

This study examined how regional differences affect elementary school students using the representative German Progress in International Reading Literacy Study (PIRLS) 2016 data (N = 3,959 fourth-grade students; M_{Age} *= 10.34 years; 49% girls; 71% from a nonimmigrant background) by combining bootstrapping, multiple imputations, principal component analysis, and the least absolute shrinkage and selection operator (LASSO). Grouping regions into rural, (sub-)urban, and metropolitan, we found that students from rural and metropolitan areas are 10.9% and 15.1% more likely, respectively, to receive an academic track recommendation than their urban counterparts. Similarly, rural and metropolitan students are 0.2 to 0.3 standard deviations more likely to enjoy school and be interested in reading than their urban counterparts. Aside from students' backgrounds and skills, many of the characteristics explaining this regional difference are structural, directly affected by policy decisions. Variables directly and indirectly influenced by policy help explain regional differences, but non-policy variables reduce regional differences in academic track recommendations the most.*

Keywords: *regional differences, school resources, teacher distribution, academic achievement, well-being, interest, elementary school, fourth grade school students, machine learning application, Progress in International Reading Literacy Study (PIRLS) 2016*

Introduction

Even slight differences in early-life conditions can produce significant social inequalities by establishing path-dependent trajectories (DiPrete & Eirich, 2006; Elder et al., 2003; Kosse et al., 2020; Organisation for Economic Co-operation and Development [OECD], 2017). This underscores the importance of ensuring the equal quality of students' institutional environments because growing up in a region with "better" characteristics improves students' educational trajectories (Chetty et al., 2016) as well as labor market outcomes (Abramitzky et al., 2021; Lavy, 2021). In particular, being part of an effective institutional learning environment during elementary school relates to educational attainment later in life (Tymms et al., 2018). Thus, whether children growing up in different regions within a country experience equal opportunities in their institutional environment along their educational pathway is a fundamental question when it comes to designing and evaluating educational policies.

Children are inevitably influenced by the regional characteristics of their surroundings—especially local educational institutions, which they cannot choose (Mayer & Koinzer, 2020).[1] Furthermore, a school's resources affect health and academic outcomes. Better-resourced schools can provide, for example, a healthier breakfast and lunch, which increases students' health (Rothbart et al., 2023). Additionally, the little residential mobility that occurs in Germany is mainly related to changes in parents' jobs, meaning that school choice is determined endogenously (Krabel & Flöther, 2014; Sánchez & Andrews, 2011). This is particularly relevant for elementary school students, who are too young to commute to a distant school and given that no school busing is provided in Germany. In addition, students in Germany must attend the elementary school nearest to their current place of residence. Hence, being born in a specific region has potentially long-term effects on children (Elder et al., 2003). Consequently, examining how regional differences affect students' school outcomes and where precisely the source of these regional variations lies is crucial. We investigated which regional differences in core educational outcomes, such as educational tracking at the end of elementary school, enjoyment of school, and reading motivation, exist and analyzed characteristics that help explain these regional differences.

Germany's education system offers a unique research context, particularly for research on early tracking, which is systematically implemented in most states starting after fourth grade, when students are, on average, 10 years old. Among Organisation for Economic Co-operation and Development (OECD) countries, only Austria has the same early tracking; the Czech Republic and Slovakia track students a year later, and Belgium and the Netherlands track students after sixth grade at age 12 years; most other countries track students only at ages 14–16 years (Woessmann, 2009). In the U.S. context, most tracking occurs with accelerated or honors courses in middle and high school, and with Advanced Placement courses in high school, so tracking around the seventh grade is relatively typical in the United States. In addition, Germany's equitable distribution of educational resources across its states ensures that studies can focus on regional differences without the confounding effects of resource variability at the state level. This combination of features makes the German education system a particularly illuminating subject for academic research.

Historically, educational outcomes were analyzed based on a rural–urban dichotomy, with U.S. students in rural areas often performing worse academically and having lower aspirations and higher dropout rates (Gershenson et al., 2017; Roscigno & Crowley, 2001; Roscigno et al., 2006; Strayhorn, 2009). Building on recent literature such ass Schmitt et al. (2018), this study adopted a three-way division—metropolitan (large cities), urban (small- and medium-sized cities), and rural. Notably, students in metropolitan areas fare better academically than those in both urban and rural regions, with urban students showing unique well-being challenges (Gross-Manos & Shimoni, 2020; Miller & Votruba-Drzal, 2013). Using data from Germany, our research delved into these regional disparities, analyzing various influential factors from students, parents, teachers, and schools.

Additionally, to find which characteristics help explain possible regional differences, we used advanced machine learning techniques to comprehensively analyze the German Progress in International Reading Literacy Study (PIRLS) 2016 dataset, allowing for a rich set of controls while mitigating overfitting and omitted-variable bias. This approach demonstrates how combining methods can robustly analyze large datasets beyond just a few selected control variables. Unfortunately, no universal agreement exists on defining and characterizing different types of regions (Rees et al., 2017; UN Statistics Division, 2017). We grouped regions into the following groups: (a) metropolitan with 100,001 inhabitants or more, (b) urban and suburban with 15,001–100,000 inhabitants (only urban thereafter), and (c) rural with at most 15,000 inhabitants.

Geographic and Educational Landscape of Germany

Many studies on regional differences are based on U.S. data, a country with high income inequality, a high degree of (human) mobility (Foster, 2017; Sánchez & Andrews, 2011), and ethnic segregation (Cutler & Glaeser, 1997; Farrell, 2008). In this study, we looked at Germany, a country with a similar level of education but less income inequality, less internal (human) mobility, and less segregation (Musterd, 2016; Sánchez & Andrews, 2011). Studying regional differences in a country such as Germany has multiple major advantages. By focusing just on one entire country, namely Germany, we are able to account for the different legislative and administrative environments within it, that is, the German federal states. This allows for key drivers of educational outcomes to be factored out more clearly in contrast to comparing variations in educational outcomes across different countries. This approach is underscored by findings of more regional variation within than between countries (Rees et al., 2017).

First, income inequality within Germany (Gini coefficient = .289) is below the median (Gini = .306) and mean (Gini = .321) of OECD countries (OECD, 2016). Countries with more prominent regional differences in education, such as the United States (e.g., Miller & Votruba-Drzal, 2013) and Israel (e.g., Gross-Manos & Shimoni, 2020), experience more severe inequality (United States: Gini = .390; Israel: Gini = .390; OECD, 2016). For comparison, the largest Gini coefficient[2] (most income inequality) in the OECD is Costa Rica, with a Gini of .487, and the smallest is Slovakia, with a Gini of .222 (OECD, 2016). Income effects, which may account for most regional variation in other countries, may not be as pronounced in Germany. Therefore, this study can potentially identify regional factors beyond income inequality.

Second, the characteristics of Germany allow for straightforward identification of regional differences in educational opportunity. Germany has various metropolitan, urban, and rural areas, even within individual federal states (the German Länder/*Bundesländer*). In Germany, the mechanisms of school choice and teacher allocation differ significantly from those in the United States and reduce the endogeneity of the regions. Specifically, in Germany, parents do not have the ability to influence which teachers are assigned to their children. Teacher assignments are mostly managed by the

Ministries of Education of the states, which allocate teachers to schools based on systemic needs rather than parental preference with few exceptions. Therefore, the notion of parents manipulating the system to secure better teachers for their children is not applicable in our setting. Additionally, the German PIRLS 2016 data (Hußmann et al., 2017) provide rich information on elementary school children, making it possible to examine systematic regional inequalities.

Third, a comprehensive meta-analysis of >200,000 job applications from nine countries found that racial inequality in the labor market in terms of job-application response rates differs between the United States and Germany and other countries (Quillian et al., 2019). The United States has a unique demographic composition, with a significant population descended from enslaved people. In Germany, the largest minority group has a Turkish ethnic background (2.5 million people, i.e., 3% of the population, according to the German Federal Statistical Office [2010]). This is rooted in the recruitment of *Gastarbeiter* ("guest workers") from Turkey by the West German government in the 1960s to expand the labor force (Kaas & Manger, 2012). Polish and Russian immigrants—the majority of whom immigrated to Germany after reunification in 1990—form the second and third largest minority ethnic groups in Germany. However, none of these immigrant ethnicities are recognized officially as minorities by the German state. The German state recognizes four national minorities: the Danish minority, the Frisian ethnic group, the German Sinti and Roma, and the Sorbian people. However, no statistics are collected on these groups. These historical differences, combined with different institutional approaches to race and hiring, have significantly impacted patterns of discrimination in each country (Quillian et al., 2019). Although both Germany and the United States exhibit racial discrimination, discrimination in Germany tends to be less severe and more related to statistical discrimination based on limited information about applicants. In contrast, the United States has a more pronounced racial gap evenly distributed across different job sectors (Quillian et al., 2019).

Finally, schools' physical infrastructure and resources are relatively homogeneous within federal states in Germany, unlike in other countries and especially unlike school funding in the United States. While in the United States the federal government tries to mitigate disparities in educational spending by supporting states and local municipalities with 8%, on average, of total expenses for elementary and secondary education (although in some states, support from the federal government makes up almost half the educational funding; Hanson, 2022), in Germany, schools are financed by the federal states (the *Bundesländer*) together with municipal institutions (usually governmental institutions associated with the European administrative Nomenclature of Territorial Units [NUTS] 3 level). Broadly speaking, personnel costs are paid by the federal state, whereas municipal institutions pay expenses for land, buildings, and materials. Exceptions are building construction expenses or (digital) equipment, which may be cofinanced by support programs from the federal states. The federal government has recently become more active with federal funding programs such as the *Digitalpakt* for digital equipment. In Germany, schools themselves have only limited opportunities to acquire additional funds (Köller et al., 2021), and local taxes play essentially no role in financing schools. Public expenses per student in elementary and secondary schools in Germany in 2020 ranged from €13,000[3] (in the state of Berlin) to €8,100 (in North Rhine–Westphalia; German Federal Statistical Office, 2021), whereas public educational expenses for elementary and secondary schools in the United States ranged in 2022 from $24,881 (in New York) to $8,041 (in Idaho; Hanson, 2022). For Germany, expenditure for elementary schools alone ranged from €12,100 (in Hamburg) to €6,300 (in North Rhine–Westphalia) per enrolled student. Differences due to school- or region-specific budget constraints therefore can be implicitly controlled for in our analysis by accounting for the federal state in which a specific school is located (Rees et al., 2017).

Important School-Related Student Outcomes: Academic Track Recommendation, Enjoyment of School, and Reading Interest

Many factors contribute to children's success in (elementary) school. Academic achievement, school-related well-being, and motivation are three of the most critical factors. *Academic achievement* refers to the actual knowledge or skills learned or acquired and can be defined as the successful completion of a level of education or learning. In elementary school, aspects of academic achievement include competencies, grades, and transition recommendations. All three are crucial for educational trajectories, especially in a school system with different academic tracks starting in secondary school. In Germany, there are essentially four tracks after elementary school: *Hauptschule* (which awards a [lower] secondary school leaving certificate, which runs until ninth or tenth grade), *Realschule* (which awards an intermediate secondary school leaving certificate, which runs until tenth grade), *Gesamtschule/Gemeinschaftschule* (a comprehensive school integrating different educational tracks within a single institution, providing a holistic approach to education for students of varying abilities and backgrounds, where all school leaving certificates can be acquired), and *Gymnasium* (which awards the highest school certificate allowing for direct university admission, also known as the *academic track*, which runs until twelfth or thirteenth grade). Transition recommendations allocate students into these four (main) tracks at an early age, usually at the end of fourth grade, at around age 10 years. These recommendations can have long-lasting effects on children

(Legewie, 2012). Although students are, in principle, able to change school tracks later, this is rarely the case (Biewen & Tapalaga, 2017). Our study emphasized the post–fourth grade track recommendation, which captures a student's cumulative performance in elementary school across multiple domains (Schwerter, Stang-Rabrig, et al., 2024). This recommendation is not determined by standardized test scores but by combining student grades, teacher evaluations, and a holistic assessment of the student's abilities. Although teachers give this recommendation to parents, its binding nature varies from state to state, with some treating it as a mere prediction of where a student might optimally thrive, called a *promotion prediction* (*Förderprognose*). It is difficult to define a ranked order of the types of secondary schools, primarily because of the *Gesamtschule/Gemeinschaftschule*, but there are significant achievement differences between students in the academic track and students in other tracks (Hübner et al., 2017; Maaz et al., 2020), and students in the academic track have historically made up the vast majority of university students (94% in 2004; Education Reporting Consortium, 2006). Therefore, our variable has potentially lifelong implication for students because it can approximate student's future level of education or a "graduation" in fourth grade.

Given the importance of fourth grade in shaping a student's educational journey in Germany (Biewen & Tapalaga, 2016, 2017; Legewie, 2012; Maaz et al., 2020), our study also addressed student well-being and motivation. During this transitional period, well-being and motivation are particularly important (Cai et al., 2020; Govorova et al., 2020; McBreen & Savage, 2021; Seligman & Csikszentmihalyi, 2000; Steinmayr et al., 2018; Wigfield et al., 2008) because they can directly impact a student's ability to adapt, engage, and excel in their recommended academic pathway. A positive sense of well-being increases a student's resilience and fosters a conducive environment for learning and personal growth (Cai et al., 2020; Govorova et al., 2020; OECD, 2017; Seligman & Csikszentmihalyi, 2000; Steinmayr et al., 2018). Meanwhile, strong motivation is a driving force that enables students to overcome challenges and remain curious and committed to their academic pursuits (McBreen & Savage, 2021; Wigfield & Eccles, 2020; Wigfield et al., 2008). As students prepare to navigate the new demands and expectations of their chosen pathways, their emotional state and intrinsic drive play a critical role in determining their subsequent academic success and satisfaction. *Subjective well-being* is a general term to describe people's subjective perception of their own well-being (Diener & Ryan, 2009). Because students' well-being also reflects an educational system's effectiveness (Seligman et al., 2009; Steinmayr et al., 2018), it is a crucial educational goal in elementary school. In this study, we operationalized students' school well-being as their enjoyment of school, a prominent aspect of school-related positive affect. Enjoyment of school captures whether students like going to school (Hascher, 2008; Kleinkorres et al., 2022). When students feel safe and supported at school, they are more likely to focus on their academic performance. Establishing nurturing relationships with peers and educators contributes positively to both academic achievement and well-being (Seligman et al., 2009).

Finally, we considered students' interest as a motivational variable (Hidi & Renninger, 2006) because it is on par with academic achievement and well-being in its importance for students (Vu et al., 2022; Wigfield et al., 2016). *Interest* is defined as the interaction among individuals' goals, cognition, and affect as well as the environment (Hidi & Renninger, 2006; Walgermo et al., 2018). Interest includes both affective and cognitive components via separate but interacting systems (Hidi & Renninger, 2006). The development of interest is related to different levels of learning and several critical educational outcomes. Attention, goal setting, achievement, the effective use of learning strategies, and the regulation of behavior play critical roles (Hidi & Renninger, 2006; Renninger & Hidi, 2020). Particularly in childhood, interest in reading is essential for students' skill development. *Reading interest* is often defined as specific values and goals that encourage a student to read (Vaknin-Nusbaum et al., 2018). For elementary school students, research shows the importance of reading interest for reading comprehension and achievement (Asadi, 2018; Vaknin-Nusbaum et al., 2018) and academic achievement in general (Cooper et al., 2014; Kirsch et al., 2002). Therefore, reading interest is a central prerequisite for reading skills (Cantrell et al., 2018; Ivey & Johnston, 2013; McBreen & Savage, 2021; Snow & Sweet, 2003; Wigfield et al., 2008) and achievement in all other school subjects (Leiss et al., 2019). Additionally, it is essential for education, work, and social participation (McBreen & Savage, 2021; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2005).

Evidence for Regional Differences in Achievement, Well-Being, and Reading Interest

Distinguishing between metropolitan, small urban, suburban, and rural regions, Miller and Votruba-Drzal (2013) examined children's language and math skills in kindergarten in the United States. They found a reverse U-shaped effect: Children in metropolitan and rural areas entered kindergarten with less advanced academic skills than children in small cities and suburban areas. The lower achievement of children in rural areas was partly explained by less advantaged home environments as well as greater use of home-based preschools. In the study by Miller and Votruba-Drzal (2013), parents in metropolitan areas were less knowledgeable about children's development in general, which accounted for their children's poor performance. A similar reversed U-shaped effect to the one described by Miller and Votruba-Drzal (2013) can be found in the work of Roscigno et al. (2006): Rural as well as urban eighth grade students are more likely to drop out of high school than their suburban counterparts. These findings are also consistent with the work of Roscigno and Crowley (2001), who showed that families from urban and rural areas were disadvantaged regarding family and school resources. These resource inequalities led to lower educational investments at the family and school levels, which explained the deficits in school completion and standardized achievement (Roscigno & Crowley, 2001).

Furthermore, studies have shown that children in rural areas report higher well-being than children in urban areas, for example, among children aged 8–12 years in Spain (González-Carrasco et al., 2019). This result is consistent with that of Rees et al. (2017). They found that 8-, 10-, and 12-year-old children in South Korea, Argentina, South Africa, and Romania have slightly higher subjective well-being in rural areas than in urban areas. However, partitioning regions in Israel into three categories, Gross-Manos and Shimoni (2020) found that elementary school children (mean age = 10 years) living in suburban areas reported significantly lower subjective well-being than those in metropolitan and rural areas. Hence, a reversed U-shaped effect is also present in this sample. Additionally, Gross-Manos and Shimoni (2020) found that children living in metropolitan areas with fewer material resources reported lower scores for school satisfaction. However, they reported higher satisfaction with their day-to-day social encounters at school.

Results regarding the effects of regional differences on students' motivation are mixed (Alejo & Piquer-Píriz, 2016; Lamb, 2012). In particular, analyses comparing students' reading interest across different regions are scarce (Cantrell & Rintamaa, 2020). For example, Cantrell et al. (2018) studied only rural students, finding an increase in students' intrinsic value of reading and out-of-school reading behavior after their transition from middle to high school and that increased reading motivation is related to reading performance. Similarly, Cantrell and Rintamaa (2020) showed that reading is vital to students' lives in rural areas. However, they did not include students outside rural areas either. Therefore, a comparison across regions is lacking. Hence, little is known about whether students have different levels of reading interest depending on regional characteristics (Cantrell et al., 2018).

To sum up, although some studies on regional differences in elementary school have concentrated on student well-being (González-Carrasco et al., 2019; Gross-Manos & Shimoni, 2020; Rees et al., 2017), to our knowledge, research examining the specific effects of regions on academic achievement and reading interest in elementary school is scarce to date.

Evidence on Systematic Regional Differences in Students' Environment

Regional differences among rural, urban, and metropolitan regions exist for children, their parents, the neighborhoods and cities children live in, and school resources such as staff, facilities, and equipment. For example, several international studies found that parents in rural areas were less emotionally accessible to their children than parents in metropolitan areas (Argentina, Italy, and the United States: Bornstein et al., 2008) and employed fewer positive parenting practices because they had less knowledge of child development and lower educational expectations (United States: Miller & Votruba-Drzal, 2013).

Another source of regional variation is the unequal supply of teachers across regions (Bastian & Fortner, 2020; Mathis, 2003; Monk, 2007). The local nature of the labor force, differential graduation rates of high school and college students, and the prevalence of hard-to-staff schools reinforce existing deficits in local teacher labor supply (Reininger, 2012). Nonsalary and job-quality factors also affect the supply of new teachers across urban, suburban, and rural schools in the United States (Bacolod, 2007). Additionally, because teachers often stay in the region where they obtained their first teaching experience (Krieg et al., 2015), teachers' low mobility plays a role. For example, low-income, low-achieving, and non-White students, particularly in urban areas in the United States, are more likely to be in classrooms with the least skilled teachers (Lankford et al., 2002). This unequal supply of teachers affects student development, classroom quality, teacher quality, and teacher–student interactions (Europe: García-Crespo et al., 2021; United States: Barrett et al., 2022; Gershenson et al., 2017; Schmitt et al., 2018). For example, a good teacher–child relationship in kindergarten improves math achievement in first grade (McCormick et al., 2013). Moreover, like highly qualified teachers, school principals also prefer to work in schools or districts that are well resourced, safe, and have relatively easy-to-serve students (Bertoni et al., 2023; Lankford et al., 2002; Loeb et al., 2010). Therefore, schools in disadvantaged regions also typically have less-experienced principals (Loeb et al., 2010), exacerbating the unequal teacher quality distribution.

Furthermore, there is an unequal distribution of childcare, prekindergarten, internet access, and high-quality school services across the United States (Edwards, 2021; European Commission, 2010; Fuller & Liang, 1996; Salemink et al., 2017; Thomä, 2023), resulting in unequal early opportunities. This proximity constraint negatively affects students' school experience and choices, particularly students from families with a lower socioeconomic status (Edwards, 2021). In addition, findings from the United States imply that most rural schools have less funding available than nonrural schools (Mathis, 2003; Monk, 2007). In addition, children in rural areas had less access to supportive educational materials at home and (in this pre–COVID-19 pandemic world) could not compensate for this shortage with help from nearby libraries because rural communities had less (or no) supply of developmental resources (United States: Miller & Votruba-Drzal, 2013). The lower accessibility of educational materials and public or private nonprofit institutions such as museums, libraries, and theaters can affect the school readiness of rural children (United States: Burchinal et al., 2014). In addition to educational resources, children in rural areas have less access to other developmental resources, such as mental health services (United States: Skinner & Slifkin, 2007).

The Present Study

Despite the crucial relevance of elementary school for children's educational paths, there are only a few studies on systematic differences by region in academic achievement in elementary school. Results on children's subjective well-being are mixed, with some evidence of higher well-being among children living in rural areas (Gross-Manos & Shimoni, 2020). Furthermore, no studies examine differences in academic achievement in elementary school and reading interest in general across regions (Cantrell et al., 2018). Similarly, there are only a few results on the origins of these regional differences besides the socioeconomic status of students' families (Miller & Votruba-Drzal, 2013). More insights into whether such regional differences exist and which characteristics are likely drivers of them are important due to the high impact of the learning environment in general as well as academic achievement, well-being, and interest in elementary school on later schooling outcomes.

Hence, this study contributes to the literature by filling several research gaps and connecting different strands of the literature concerning the following research questions. First, we analyzed whether there are structural regional inequalities in transition recommendations for the academic track (*Gymnasium*; hereafter *academic track recommendations*) as a measure of students' academic achievement, enjoyment of school as a measure of students' well-being, and reading interest as a measure of students' motivation among elementary school children in Germany (Research Question 1 [RQ1]). With the second research question, we wanted to go further and examined whether school, teacher, family, and student information from the rich PIRLS 2016 data could explain these regional differences. This is consistent with the call for more integrative research that does not focus on a small group of variables but takes an overarching approach (Byrnes & Miller, 2007). Therefore, we included variables in the regression for RQ1 that had been selected as important predictors of regions[4] (RQ2a) as well as the outcome variables (RQ2b) and to see whether the regional differences remained after controlling for the additional variables. With this two-step procedure, we follow the post–double selection model proposed by Belloni et al. (2014). Lastly, we compared the selected variables for the regions (RQ2a) and the outcome variables (RQ2b) to see which regional characteristics are essential for both the regions and the outcomes and therefore likely reasons for the regional differences in students' outcomes (RQ3).

Because regional divisions are a product of the social and economic constellation of a particular society, they are unlikely to be the same or similar in different countries around the world. As a result, it is unclear whether results on regional differences from one country can be extrapolated to another. Based on existing research for kindergarten and secondary school, we expected lower academic track recommendations in students from metropolitan and rural areas, similar to the effects described in the literature (Miller & Votruba-Drzal, 2013; Roscigno et al., 2006). Following Gross-Manos and Shimoni (2020), we also expected higher school well-being in metropolitan and rural areas than in urban areas. Because prior studies had failed to include students from different types of regions, our study is the first to document regional differences in reading interest. Additionally, we expected to find regional differences mostly to the detriment of rural areas in terms of socioeconomic status, parental involvement in student education, teacher quality, and school resources such as staff, facilities, and equipment, as reported by school principals. Which characteristics explain both the regions and the outcome variables are exploratory. However, following the literature, we expected teacher and resource variation to be particularly relevant.

Methods

Participants

The sample contained 3,959 fourth grade students, of which 49.18% were girls, and the mean age was 10.34 years (SD = 0.52). Only 5.17% of the students were born outside Germany, and 16.97% of the students had two immigrant parents, whereas 28.78% had at least one immigrant parent. Furthermore, 17.20% of the students were being raised by a single parent, 8.59% had at least one stepparent, and 2.63% were adopted. The mean highest international socioeconomic index (HISEI) was 53.96 (SD = 20.24) and therefore slightly above the OECD Programme for International Student Assessment (PISA) 2018 results for Germany, which yielded a mean of 51.8.

Data

This study was based on international and national data from the PIRLS 2016 (Hußmann et al., 2017) from the International Association for the Evaluation of Educational Achievement and the Institute for Educational Quality Improvement. PIRLS 2016 is a representative large-scale

TABLE 1
Summary statistics for dependent variables and regional dummy variables

Statistic	*N*	Mean	SD	Min	Max
Outcome					
Academic track recommendation	3,856	0.36	0.48	0	1
Enjoyment of school	3,493	3.30	0.66	1	4
Reading interest	3,448	3.02	0.73	1	4
Regional dummy variables					
Rural region	3,603	0.42	0.49	0	1
Suburban or urban region	3,603	0.24	0.42	0	1
Metropolitan region	3,603	0.34	0.47	0	1

assessment study with ~4,000 elementary school children. We included the student, parent, teacher, and school principal data because they all contribute to the students' social environment (García-Crespo et al., 2021; Lankford et al., 2002; Loeb et al., 2010). Thus, we ended up with 831 variables, including dummy variables for categorical data. An overview of the variables included is given in Table 1 and online Appendix Table A1. Thus, to handle this large number of variables—our dataset is very large compared with the number of observations in our dataset ($N = 3959$)—we will use dimension-reduction methods, which we describe below.

Dependent Variables. In PIRLS 2016, schools provided the elementary school teacher's recommended secondary school track for each student. Because there are many options in the German school system after elementary school, the dataset allows for 14 options. Because these options are not ordinal and cannot all be ranked, we facilitated interpretation by dichotomizing the variable according to whether students were recommended for the academic (highest) track (*Gymnasium*) or not. Thirty-six percent of students received this recommendation in the sample.

We operationalized students' subjective well-being (Fiorilli et al., 2017) as a latent variable measured by five items focusing on students' *enjoyment of school* as one central aspect of students' school-related affect. Due to the importance of the school in which a student is enrolled in the respective student's life, especially in elementary school (Eccles & Roeser, 2009), enjoyment of school is a close measure of students' general subjective well-being. Students rated their agreement on a 4-point Likert scale. Two examples are "I like going to school" and "At this school I feel like I belong" (Eccles & Roeser, 2009). The enjoyment-of-school variable had acceptable reliability (Cronbach's $\alpha = .79$).

We used students' general interest in reading activities as an indicator for motivation in the study. Here, interest is conceptualized as an affective state reflecting students' subjective learning experiences (Walgermo et al., 2018). The items in PIRLS 2016 appropriately address reading not only in the classroom but also in the out-of-school context. Three example items of the eight are "I find reading boring," "I want to have more time to read," and "I enjoy reading." We focused on reading interest because higher interest entails greater student motivation and engagement (Walgermo et al., 2018), especially for younger children (Wigfield et al., 2004). The reading interest scores had adequate reliability (Cronbach's $\alpha = .86$). Summary statistics for all three outcome variables can be found in Table 1.

Definition of Regions. In PIRLS 2016, school principals must answer how many inhabitants live in the municipality in which the school is located as follows: (a) up to 3,000, (b) 3,001–15,000, (c) 15,001–30,000, (d) 30,001–50,000, (e) 50,001–100,000, (f) 100,001–500,000, and (g) $>$500,000. Unfortunately, no global agreement exists on defining and characterizing different types of regions (Rees et al., 2017; UN Statistics Division, 2017). For instance, in Greece, the limit for a rural municipality is 10,000 inhabitants; in Norway, it is 200; in Japan, it is 50,000; and in Argentina and Bolivia, it is 2,000 (Rees et al., 2017). For our analysis, we grouped regions as follows: (a) metropolitan with 100,001 inhabitants or more, (b) urban and suburban with 15,001–100,000 inhabitants (only urban thereafter), and (c) rural with at most 15,000 inhabitants.[5] Therefore, we follow the literature separating regions into at least three different groups. We are left with 82 schools in metropolitan areas, 47 schools in urban areas, and 60 schools in rural areas. Summary statistics for the three regions on the student level are shown in Table 1.

Additional Variables. The student questionnaire included 58 questions (with many subquestions) on demographic information (e.g., gender and age), living and learning conditions in the family (e.g., social and cultural background), and school learning (e.g., individual learning conditions and school-related attitudes and perceptions). Because some of these variables were nominal variables that we recoded as a set of dichotomous dummy variables, we ended up with 135 variables, including dummy variables for categorical data. An overview of the included variables is given in Table 2 and online Appendix Table A1.

TABLE 2

Variables grouped by the taxonomy of Progress in International Reading Literacy Study (PIRLS) 2016

Variable groups	No. of survey questions	Examples of variables per group	No. of principal components
Student data			
Demographic data	3	Gender, year of birth	2
Social background	7	Persons in household, possession of books and other goods indicative of wealth, income	2
Cultural background	4	Country of origin of the student and parents, German as the native language	1
Parental support behavior	2	Parents' interest in the child's activities ("My parents always know where I go after school")	1
Extracurricular activities	1	Activities outside school	2
Individual learning conditions	6	Absenteeism at school, library usage, extracurricular reading activities, reading self-concept	1
Subject-specific teaching in the classroom	14	Reading activities in class, perceived classroom management, perceived support from the teacher	2
School-related attitudes and perceptions	1	Sense of safety at school	1
Tutoring	1	Use of private tutoring	1
Participation in all-day activities	1	Participation in all-day school activities	1
Computer and internet use	3	Computer use: location, frequency, and duration	2
Student-specific information from the student participation list	13	Grades, learning and social behavior	2
Skills	2	Cognitive skills test, receives special education support	2
Parent data			
Reference person	1	Person filling in the survey	2
Cultural background	26	Germany as the child's country of birth, language(s) in early childhood, parents' cultural practices, educational expectations	2
Social background	22	Reading habits of parents, number of books at home, parental education background and occupation, economic and cultural capital	3
Digital media at home	2	Number of digital devices in the household	2
Preschool years	4	Learning to read: preparatory activities at home, early-childhood education, child's reading skills at school entry	2
Elementary school years	5	Parental support activities at home, parents' interest in the child's friendships	3
Parents and schools	4	Parents' satisfaction with the school and child's school performance, contact between home and school	3

(continued)

TABLE 2 **(CONTINUED)**

Variable groups	No. of survey questions	Examples of variables per group	No. of principal components
Teacher data			
Demographic and educational biographical information	7	Gender, age, highest educational attainment, professional degree	1
Professional education and attitudes toward the teaching profession	8	Teaching experience, university major	3
School structure and organizational characteristics	6	School atmosphere, communication with fellow teachers, willingness to innovate among the teaching staff	3
Structure, process, and organizational characteristics of subject teaching	17	Classroom disturbances, constructivist beliefs, cognitive activation	5
Spelling instruction	3	Spelling lessons (time to practice, materials, and methods)	3
Use of phonics tables in teaching	5	Use of phonics tables (time, form, frequency, and own opinion)	1
Collaborative teaching	8	Students with and without special educational needs (preparation, adapted instruction)	1
Resources	7	Time for reading instruction/reading activities, computer use in reading instruction, classroom library, reading specialist(s) on staff	3
Design of reading instruction	23	Differentiation in reading lessons, text types in reading instruction, methods in reading instruction, teaching reading strategies, support for weak and strong readers	3
School principal data			
School spirit	3	School spirit among teachers, parents, and students	1
Structure and organizational characteristics of school and teaching	7	Number of teaching days per year/week, performance orientation of the school, classroom management: organization of learning groups	3
Resources and equipment	11	Free meals at the school,[6] homework supervision, number of computers/e-books available, problems with the school's personnel and material equipment, all-day school	3
Additional support services at the school	1	School support and care services	3
Measures for promoting reading	3	Systematic development of reading skills/strategies, measures to promote reading	3
Special education	14	*Inklusionsschule*, e.g., schools that include both students with and without special needs (implementation, year, number of students, specialized support without diagnoses, type of support needs, persons involved, and framework)	2
Immigration	5	First-generation immigrants and refugees (number of students, organization, challenges)	2
School profile and profile classes	6	School profile (type, goals), profile classes (type and goals)	3
Information about students	7	Number of students, social backgrounds, German as native language	3
Information about teachers	6	Problematic behavior by teachers, age composition, teacher cooperation	4
Information about school principal	12	Gender, experience, areas of activity of school administration	2

Note. The complete list of variables is shown in online Appendix Table A1 and follows the PIRLS 2016 taxonomy. The number of survey questions does not equal the number of items because most questions include several items. The number of principal components refers to the number of principal components we included per group in the variable selection. This is further explained below.

Second, the parent questionnaire comprised 64 questions whose answers were encoded in 241 variables. It included information on living and learning conditions in the family as well as on the child's previous learning trajectory (e.g., educational biographical information). Also, student- and parent-specific information (e.g., education background and occupational status) were included.

Third, the teacher questionnaire comprised 84 questions—of which several involved subquestions or follow-up questions. The questions covered the following topics: demographic and educational biographical information for the teacher, professional training and attitudes toward the teaching profession, working conditions at the school, characteristics of subject instruction, and specifics of reading instruction. Teacher responses to questions about themselves, the school environment, their views on the teaching profession, and their teaching style were encoded in 204 variables.

Last but not least, the school principals were asked to complete 75 questions involving structural information about the school and the people therein. This information included, for example, their professional experience, the school's enrollment scheme, available resources, instructional time, and the school's emphasis on academic success, discipline, and safety. Their responses were encoded in 251 variables.

In general, we included all variables we could generate from the respective items. We only excluded items if they were conditional on prior responses (filter questions) or if the variance was very low. For example, just two of 208 schools offered differentiation and support services for students as part of their instructional development, which is why information about these services was left out of the analysis. To facilitate interpretation of the variable scales, we changed the integer codes for some variables so that a higher integer represented a better/greater outcome or higher frequency. Additionally, we calculated scores for blocks of questions presumably related to a single latent variable. In so doing, we checked Cronbach's α. If Cronbach's α was below .70, we used the items without any further processing. Generally, we excluded observations with only missing information and constant variables without variation (minimum = maximum). We also combined student and parent data into a student-level dataset, teacher and principal data into a school-level dataset, and all together into an all-in-one dataset. Organizing the variables according to whether they capture student or school characteristics was not possible because school characteristics measured on the student level also include some additional student information. For example, if measured by asking students, the variable *teacher support* will include information about how individual students perceive that support (even if teachers manage to provide the same support to all students).

Statistical Analysis. Three types of statistical analysis were conducted to answer the research questions, as shown in Figure 1. First, we ran a baseline (OLS) regression to examine whether regional differences in our three outcome variables exist in Germany. We conducted all regressions on each school-bootstrapped sample; $B = 300$ for each dataset. We followed Rubin (1987, pp. 87ff) by pooling the estimated results. To assess the statistical significance of the pooled bootstrapped estimates, we calculated school-clustered and heteroskedastic-robust standard errors including survey weights for the PIRLS 2016 data using the survey package in R (Lumley, 2010, 2021; see online Appendix for more information). Next, we used the variable-selection method least absolute shrinkage and selection operator (LASSO; Hastie et al., 2009) to select important principal components explaining each of the three regions and each of the three outcome variables. We conducted this step once only using the student-level dataset (student and parent data), once for the school-level dataset (teacher and school principal data), and once for all data.

After the selection(s), we then included the selected principal components in a post-LASSO OLS regression in a two-step procedure. First, we included only the selected principal components for the regions and, afterwards, for both the regions and the outcomes. When estimating the standard errors of the parameter estimates in the fourth step, the specifics of the PIRLS 2016 design and the implications of the preceding steps need to be accounted for. Lastly, we compared the selected principal components for the regions and outcomes to evaluate which components influence both and are hence a possible channel through which regions affect student outcomes in elementary school. The following subsections provide the details and rationale for using the methods outlined.

OLS Regression for Baseline Results

OLS regressions were performed on each bootstrapped and imputed dataset to estimate the relationships between the outcome variables *academic track recommendation*, *enjoyment of school*, and *reading interest* and the region factors for rural and metropolitan areas. The linear model estimated with OLS can be expressed as

$$y_i = \alpha + \rho_1 rural_i + \rho_2 metropolitan_i + \gamma' \times c_i + \varepsilon_i \tag{1}$$

where the index i denotes the level of analysis and refers either to the ith individual or school; α captures the intercept for the dataset at hand; ε_i is the idiosyncratic error term; the outcome y_i is a placeholder for the outcome variables *academic track recommendation*, *enjoyment of school*, and *reading interest*; and the coefficients ρ_1 and ρ_2 capture the respective regional effects of a rural (<15,000 inhabitants) and metropolitan (>100,000 inhabitants) area, respectively, relative to an urban area. The coefficients therefore capture possible homogeneous regional effects, even though regions are influenced by the unique social and economic constructs

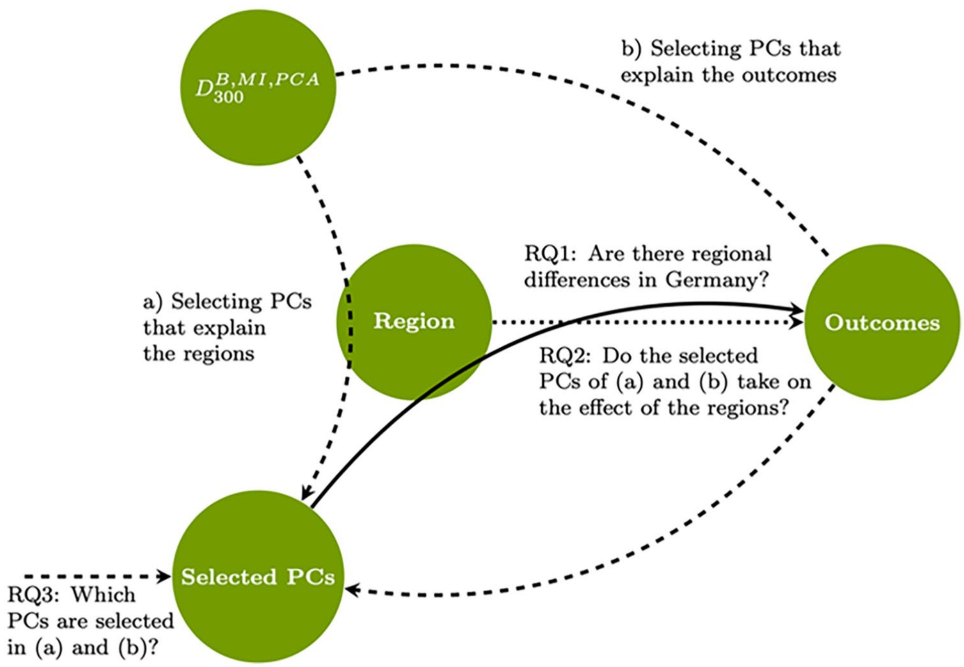

FIGURE 1. *Graphical overview of statistical analysis.*

of each society, and it is a priori unclear if there are homogeneous regional effects. The vector c_i is a placeholder for additional control variables, and γ denotes the vector of respective coefficients. In the first step, for our baseline regressions, this vector only includes fixed effects for the German federal states to absorb resource differences, specific school legislation, and specific structures described in the section "Geographic and Educational Landscapes of Germany."[7] Including state fixed effects is an attempt to account for these general differences within Germany. In further regressions, these control variables will additionally contain selected principal components from the thematic groups, as described below. This step is important to see whether the regional differences can be explained and which control variables explain them. Thus, including additional control variables is an attempt to discern further which factors may help explain the regional disparities observed in our outcome measures. This is important because the regions themselves cannot have a causal effect on students' school outcomes but rather the characteristics of these regions. Therefore, the regional classifications are an important way to capture (e.g.) local socioeconomic conditions. If the coefficients of the regions become insignificant after the inclusion of these additional variables, our results underscore this point.

Data Preparation and Dimension Reduction of Independent Variables. The statistical methods used in this paper on the path identifying the effects of regional differences address common issues researchers face when working with the PIRLS 2016 dataset. The use of these methods is somewhat novel in the context of educational research and particularly in the context of the PIRLS 2016 dataset. This dataset encompasses multilevel data with missing values. Multiple imputation is the best solution to include as much information as possible (i.e., observations and variables). However, to apply feature selection to examine which characteristics explain the regional differences, we cannot simply impute the data. The inference should account for the variance in missingness and the sample variance simultaneously. Therefore, we bootstrapped before imputing missing values to be able to assess the stability of the feature selection (Long & Johnson, 2015). The detailed data-preparation steps are shown in Figure 2 and explained below.

First, we merged the four datasets for students, parents, teachers, and school principals into one merged dataset at the student level (denoted X in Figure 2). Second, we used a block bootstrap at the school level to keep entire schools together. For imputation, we included school dummy variables, aggregated school-level information, and maintained student-level information ($A = \{X_{[S],k}^{(b)}\}_{k=1}^{300}$). We used the missRanger package for imputations for both S and A because it conducts a (fast) missing-value imputation by chained random forests. We used predictive mean matching to impute only existing realizations of the respective variables, preserving the empirical distribution (van Buuren, 2018, chap. 3.4). Furthermore, we set a maximum tree depth of six and the split rule extratrees (Geurts et al., 2006), 100 trees (Stekhoven & Bühlmann, 2012), and a sample fraction of 0.1 (Stekhoven & Bühlmann, 2012). We relied on missRanger (Mayer, 2019)

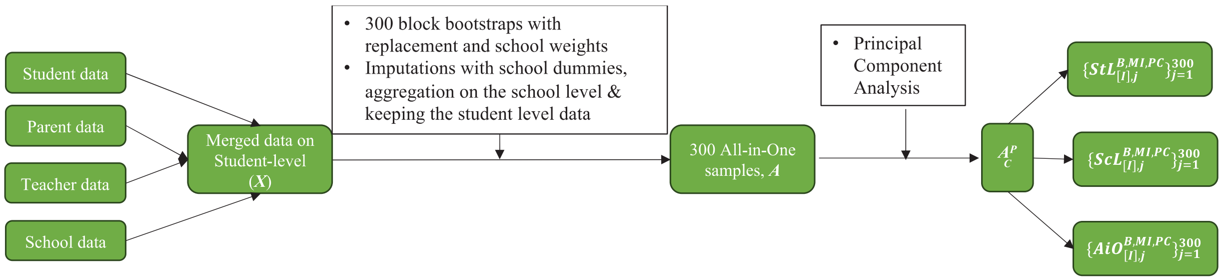

FIGURE 2. *Graphical overview of data preparation.*
StL, student level; ScL, school level; AiO, all-in-one, meaning it includes student- and school-level information; B, bootstrap; MI, multiple imputation; PC, principal components.
Note. The three datasets StL, ScL, and AiO were all resampled with replacement, missing values were imputed 300 times, and the principal components were generated.

for our mixed and high-dimensional data because tree-based methods are recommended over mice with predictive mean matching in part given that mice with predictive mean matching can be overly conservative (Akande et al., 2017; Murray, 2018; Schwerter, Gurtskaia, et al., 2024). Additionally, Madley-Dowd et al. (2019) showed that using the tree-based method reduces bias even in the presence of a large proportion of missing values.

Third, we applied principal component analysis (PCA) to streamline the analysis and reduce the dimensionality, similar to Hanushek et al. (2022), who reduced the dimensionality of socioeconomic status with the help of PCA. According to the taxonomy from the PIRLS documentation for S and A, a group of variables is represented by a relatively low number of components. For example, there are several variables and questions in the PIRLS 2016 survey that hold information about a student's cultural background: country of origin of the father, country of origin of the mother, country of origin of the student, and when the student began to learn German. We calculated the principal components using the psych package in R (Revelle, 2020). Then, by plotting the contributions of each principal component to the overall variation exhibited, we found the point of inflection in the plot and decided accordingly how many components to consider for each group of variables. In total, we formed 40 such groups (of which 13 are related to variables from the student survey, 7 to the parent survey, 9 to the teacher survey, and 11 to the school principals survey). The taxonomy of the PIRLS 2016 variables as well as the number of principal components per group can be found in Table 2. In total, we generated 89 principal components; for most groups, we included one or two principal components, and the maximum across all groups is 5 (for structure, process, and organizational characteristics of subject teaching).

Fourth, because we wanted to run the feature selection for student- and school-level data separately, we then generated datasets that only contained the principal components from the student-level dataset ($\{StL_{[I],k}^{B,MI,PC}\}_{k=1}^{300}$) and school-level dataset ($\{ScL_{[I],k}^{B,MI,PC}\}_{k=1}^{300}$). For the student-level data we had 37 principal components, whereas for the school-level data we had 52 principal components. Adding them to the all-in-one data ($\{AiO_{[I],k}^{B,MI,PC}\}_{k=1}^{300}$) gives 89 principal components from which LASSO could select. Important to note is that the number of students per imputed dataset varies because the number of students within a school (chosen by the block bootstrap) varies. Although we have 3,959 students in our data, the imputed datasets have, on average, 3,565 students.

Variable Selection of Essential Independent Variables

As a second step after the dimension reduction of the PCA, we used LASSO regressions to select essential principal components that explain membership in the metropolitan, urban, and rural area, respectively, as well as our outcome variables *academic track recommendation*, *enjoyment of school*, and *reading interest* (denoted by z_i in the following regressions), following the post-double-selection procedure by Belloni et al. (2014). Selecting variables for both the outcomes and the regional variables was important for the two-step approach in the post-double-selection method: selecting variables that explained (a) the outcomes and (b) the variables of interest. Then we could include principal components that explain the regions (c_i^r) and variables, explaining both the region and the respective outcome variables (c_i^{ry}) as control variables in the post-selection regression. Even though we had already reduced the dimensions via PCA, we still needed to use variable selection via LASSO because we were still left with a very high number of principal components, meaning that there was still a possibility of overfitting the data with our model. Given the high-dimensional nature of our dataset and the multifaceted nature of regional differences, our approach was deliberately chosen to capture a comprehensive set of determinants of regional differences. Further, the additional selection step helped us to compare the sets to show which were selected for the regions and the respective outcome variables. Additionally, we included information about the school size and class size so that we could distinguish between pure size differences that might be present between regions and effects of the variables selected via LASSO. The remainder of this section describes our labeling method in greater detail.

We chose the LASSO approach to include a sufficient set of additional control variables c_i without overfitting. In the LASSO approach, the estimated model is extended by the LASSO penalty term $\sum_j \left|\beta_j\right|$, that is, the sum over the absolute values of the coefficient vector β, which is weighted with the coefficient λ. In this study, the estimated LASSO regression can be represented as

$$z_i = \alpha + \beta' \boldsymbol{x_i} + \varepsilon_i \tag{2}$$

where β minimizes

$$\frac{1}{2n} \sum_i z_i - \alpha - \beta' \boldsymbol{x}_i)^2 + \lambda \sum_j \left|\beta_j\right| \tag{3}$$

where the index i again represents either the ith individual or school. The index j marks the jth variable in the vector of explanatory variables in the model and thus also the jth coefficient, and z_i includes the outcome variables included in y_i as well as the regional variables $rural_i$, $metropolitan_i$, and $urban_i$, the baseline for the regional variables. Including both the outcomes and the regional variables is important for the two-step approach in the post-double-selection method. λ is the penalty weight. Including the penalty term excludes irrelevant variables by setting some β_j to zero, that is, when the respective variable x_{ij} does not add sufficient explanatory power for the outcome z_i. Note that the vector x_i includes the intercept and the calculated principal components. Again, ε_i is the student- or school-specific component unexplained by the model. Although some variables in z_i are binary, we stick with the linear probability model (also for the OLS [post-selection] regressions). This is motivated by the rich literature favoring the linear probability model (Beck, 2019; Chatla & Shmueli, 2013; Goldfarb & Tucker, 2011; Gomila, 2021; Hellevik, 2009; Horrace & Oaxaca, 2006; Lee & Hosanagar, 2021), particularly for models with a high number of fixed effects (Caudill, 1988), because we are not interested in predicted probabilities, and the coefficients still can be validly interpreted as approximate partial effects of a conditional expected value function. In addition, because the outcomes *enjoyment of school* and *reading interest* are mean scores, we chose to analyze them using linear regression rather than ordered logit models. In addition, using a linear probability model for the binary outcome of *academic career recommendation* improves the comparability of our estimation results across outcomes. Furthermore, a linear probability model may be disadvantageous only when probabilities close to zero or one are plausible, but in our case the advantages of the linear probability model in terms of interpretability outweigh this (Aldrich & Nelson, 1984). We ran the LASSO for all the student, parent, teacher, and school principal data separately. Additionally, we included the school and fourth grade class size for each data point to control for general differences due to school size.

To determine the hyperparameter λ for each sample, we conducted the LASSO regression by cross-validation on 999 folds of each of the bootstrapped 300 samples. Thus, for each sample, a different list of variables was selected, and we thus had 300 differing lists of selected principal components. We deem a variable selected if 80% of the lists contain the principal component.

Post-Selection Regression Models

The variable selection using LASSO determined which variables were important explanatory variables for the dependent variable and the regions (z_i). In the subsequent step, the post-LASSO step, we include these variables in Equation (2) as additional control variables via c_i. The principal components selected as essential in explaining the regional variation are relevant for answering RQ2. The selected variables effectively characterize regional profiles and help explain differences in academic track recommendations, school enjoyment, and reading interest. We denote the vector that contains these variables as c_i^r. When discussing the results, it will be interesting to examine whether the effect of the regions ρ_1 and ρ_2 remains statistically significant when controlling for c_i^r. Potentially, suppose that the control variables could explain the variation in the outcome variable better than the regional regressor, in a kind of horse race between the selected control variables and regional regressors. In such a case, parameter γ would pick the better fit and render the regional regressor insignificant. In this post-LASSO regression setup, the regional regressors acted as a (partial) mediator, and the specific characteristics of the regional differences modeled through the control variables provided more specific explanatory power.

We further used the suggested post–double selection in Belloni et al. (2014) and used the LASSO to select essential features not only for our regional variables but also for the outcome variables (*academic track recommendation*, *school well-being*, and *reading interest*). We denoted a vector containing a collection of these selected variables as c_i^{ry}. In this vector, the control variables that explained the regional differences c_i^r were extended to include selected control variables for the outcome variables. Note that variables may have explanatory power for the categorization of regions as well as for the outcome variables. Lastly, to address RQ3, we compared whether specific variables were selected for both the regions and (at least one) of the outcome variables. This allowed us to find the principal components through which the region influenced students' outcomes.

Results

Baseline Regional Differences

Table 3 in the first column of panels A, B, and C (model 1) shows OLS regression results following Equation (1) and answering our RQ1, including only the regional factors for rural (<15,000 inhabitants) and metropolitan regions (>100,000 inhabitants) as well as fixed effects for the states to account for general differences within Germany, such as resource differences and the like between the states. Thus, the baseline is urban regions. Our findings reveal a U-shaped regional effect on academic track recommendations. Specifically, students in both metropolitan and rural areas are significantly more likely to receive such recommendations than their urban counterparts. The probability among students in rural regions is 10.9% higher at the 5% significance level and rises to 15.1% higher for the metropolitan region at the 1% significance level (Table 3, panel A, model 1). For enjoyment of school, we found a similar U shape; that is, students in rural regions had a higher enjoyment of school by .302 of a standard deviation at the 1% significance level, and students in metropolitan regions had a higher enjoyment of school of .239 of a standard deviation at the 10% significance level (Table 3, panel B, model 1). Lastly, for reading interest, again, both regional coefficients are positive (.21 and .253) and marginally significant at the 10% significance level (Table 3, panel C, model 1).

Post-Double-Selection Regression

Next, we included the selected principal components from all categories at the student level (including student and parent survey information) and school level (including teacher survey and school principal survey information). This allowed the variables to be identified that captured the effects of the regional differences. For RQ2a, we first included only principal components selected to describe the regions (c_i^r, Table 3, model 2 for student level, model 4 for school level, and model 6 for all-in-one). We then included principal components that had been selected to describe the outcomes (c_i^{ry}, Table 3, model 3 for student level, model 6 for school level, and model 7 for all-in-one, RQ2b). Regarding our first outcome, *academic track recommendation*, model 2 in Table 3, panel A, shows that including student-level data explaining the regions decreased the coefficient of the rural region to close to zero (0.007) and reduced the coefficient of the metropolitan region by about two fifths to 0.065; both coefficients are no longer significant. Hence, the student-level data could explain some of the differences between rural and urban areas as well as between metropolitan and urban areas. Including principal components selected for the academic track recommendation, the rural region coefficient increased slightly to 0.012 and the metropolitan region decreased further to 0.05, but both remained insignificant. Comparing models 2 and 3 for the student level, we see that 18 vs 29 principal components were included and that the summary statistics R^2, R^2_{adj}, and *BIC* all favored model 3 over model 2. Next, we included the school-level principal components in models 4 and 5, first for the regions and then for the regions and the respective outcomes. Both models were identical, meaning the same variables were selected for the regions and academic track recommendations. The coefficients were about halved and nonsignificant. The summary statistics, however, showed that the school-level data explained less than the student-level data. Letting the LASSO select from both student- and school-level principal components and including them in the post-LASSO regressions in models 6 and 7 yielded very similar reduced and statistically insignificant coefficients as before while including 69 and 78 additional principal components compared with model 1. In particular, the Bayesian information criterion showed that the additional inclusion of the school-level data did not improve the general model fit compared with including only the student-level data. Thus, it seems that the school-level data introduced more noise than additional explanatory power.

The story is the same for the outcomes enjoyment of school and reading interest: Including selected student-level principal components reduced the coefficients and improved the general model fit. It seems that including additional school-level information added noise. Only for reading interest, the coefficient for rural regions flipped the sign, and students seemingly had a lower reading interest in the post-LASSO regressions. For model 3 only, the coefficient was marginally significant. It was significant for all other cases.[8]

Selection of PCs

In Table 4 we present the selection results for the principal components included in models 6 and 7 in Table 3, which form the union *p* between the group of principal components selected for profiling the three regions and the group of principal components predicting an outcome variable. This highlights the critical information explaining the regional variation in the outcome variables and thus addresses our RQ3. The following text only highlights principal components that were selected both for one of the regions and at least one of the outcomes.

Selected Student-Level Information. Student Survey Data. From the student survey data, student-, parent-, and teacher-specific information was selected. The first two principal components from the set of student-specific information from the student participation list were essential for profiling the regions as well as for academic track recommendation and enjoyment of school (only the first principal component). This group included information on student characteristics such as German and math grades as well as student

TABLE 3

Ordinary-least-squares and post-LASSO regression results

	OLS baseline	Post-LASSO					
		Student level		School level		All in one	
Factor	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Panel A	Dependent variable: *academic track recommendation*						
Rural region	0.109*	0.007	0.012	0.054	0.054	0.021	0.018
	(0.045)	(0.039)	(0.039)	(0.070)	(0.070)	(0.063)	(0.063)
Metropolitan region	0.151**	0.065	0.050	0.074	0.074	0.058	0.059
	(0.053)	(0.041)	(0.037)	(0.072)	(0.072)	(0.060)	(0.059)
c_i^r	No	Yes	Yes	Yes	Yes	Yes	Yes
c_i^{ry}	No	No	Yes	No	Yes	No	Yes
Avg. observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608
R^2	0.093	0.444	0.484	0.190	0.190	0.520	0.551
R^2_{adj}	0.089	0.439	0.479	0.174	0.174	0.509	0.539
BIC (null)	−204.5	−1,838.9	−2,060.2	−162.2	−162.2	−1,889.0	−2,050.5
Variables	18	29	34	72	72	87	96
Panel B	Dependent variable: *enjoyment of school*						
Rural region	0.302**	0.066	−0.009	0.064	0.064	−0.032	0.001
	(0.107)	(0.091)	(0.075)	(0.163)	(0.163)	(0.122)	(0.106)
Metropolitan region	0.239***	0.056	0.001	0.046	0.046	0.030	0.025
	(0.131)	(0.104)	(0.086)	(0.154)	(0.154)	(0.126)	(0.116)
c_i^r	No	Yes	Yes	Yes	Yes	Yes	Yes
c_i^{ry}	No	No	Yes	No	Yes	No	Yes
Avg. observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608
R^2	0.035	0.235	0.399	0.141	0.141	0.351	0.438
R^2_{adj}	0.030	0.230	0.393	0.124	0.124	0.335	0.423
BIC (null)	13.5	−718.0	−1,514.9	42.2	42.2	−821.7	−1,286.5
Variables	18	29	35	72	72	87	92
Panel C	Dependent variable: *reading interest*						
Rural region	0.210***	−0.075	−0.117***	−0.031	−0.031	−0.075	−0.051
	(0.109)	(0.083)	(0.060)	(0.164)	(0.164)	(0.124)	(0.092)
Metropolitan region	0.253***	0.038	−0.008	0.036	0.036	0.024	0.048
	(0.138)	(0.096)	(0.071)	(0.156)	(0.156)	(0.122)	(0.095)
c_i^r	No	Yes	Yes	Yes	Yes	Yes	Yes
c_i^{ry}	No	No	Yes	No	Yes	No	Yes
Avg. observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608
R^2	0.024	0.264	0.528	0.148	0.148	0.354	0.549
R^2_{adj}	0.019	0.258	0.524	0.131	0.131	0.339	0.537
BIC (null)	53.9	−851.7	−2,380.5	14.5	14.5	−839.5	−2,069.1
Variables	18	29	34	72	72	87	91

OLS, ordinary least squares; LASSO, least absolute shrinkage and selection operator; BIC, Bayesian information criterion

Notes. The table shows the OLS and post-LASSO regression results for the three outcome variables: *academic track recommendation* (panel A), *enjoyment of school* (panel B), and *reading interest* (panel C) using the school bootstrap sample. That is, we resampled the schools before imputing missing data, resulting in, on average, 3,608 students (avg. observations) across the imputed datasets. First, model 1 presents the OLS (baseline) regression results including only the state fixed effects as additional variables. Next, the table shows post-LASSO regressions at the student level, in which we have included selected principal components for regions (model 2) and for regions and the respective outcomes (model 3). The same is done for school-level principal components (models 4 and 5) and selected student- and school-level principal components (all-in-one, models 6 and 7). Regression results without state fixed effects can be found in Online Appendix: https://osf.io/dyj3h/?view_only=0de4096d6090400694b99a2d0cdfaeb7.[8]

*$p < .05$; **$p < .01$; ***$p < .1$.

TABLE 4

Principal component selection for the all-in-one data

Variable group	PCs	Examples	Region	ATR	EoS	RI
Size						
School size	1	Number of students in the school	X	—	—	—
Fourth grade class size	1	Number of students in the class	X	—	—	—
Student level						
Student survey						
Demographic data	2	Gender, year of birth	—	X	—	—
Social background	2	Persons in household, possession of books and other goods indicative of wealth, income	—	X	—	—
Cultural background	1	Country of origin of the student and parents, German as the native language	—	—	—	—
Parental support behavior	1	Parents' interest in the child's activities ("My parents always know where I go after school")	X	—	—	X
Extracurricular activities	2	Activities outside school	—	X	X	X
Individual learning conditions	1	Absenteeism at school, library usage, extracurricular reading activities, reading self-concept	—	—	X	X
Subject-specific teaching in the classroom	2	Reading activities in class, perceived class management, perceived support from the teacher	X	—	X	X
School-related attitudes and perceptions	1	Sense of safety at school	—	—	X	—
Tutoring	1	Private tutoring	—	—	—	—
Participation in all-day activities	1	Participation in all-day school activities	X	—	—	—
Computer and internet use	2	Computer use: location, frequency, and duration	X	—	X	—
Student-specific information from the student participation list	2	Grades, learning and social behavior	X	X	X	—
Skills	2	Cognitive skills test, receives special education support	X	X	—	—
Parent survey						
Reference person	1	Person filling in the survey	X	—	X	—
Cultural background	2	Germany as the child's country of birth, language(s) in early childhood, parents' cultural practices, educational expectations	X	—	—	—
Social background	3	Reading habits of parents, number of books at home, parental education background and occupation, economic and cultural capital	X	X	—	—
Digital media at home	2	Number of digital devices in the household	X	—	—	—
Preschool years	2	Learning to read: preparatory activities at home, early-childhood education, child's reading skills at school entry	—	X	—	—
Elementary school years	3	Parental support activities at home, parents' interest in the child's friendships	X	X	—	—
Parents and school	1	Parents' satisfaction with the school and child's school performance, contact between home and school	—	X	X	—

(continued)

TABLE 4 **(CONTINUED)**

Variable group	PCs	Examples	Region	ATR	EoS	RI
School level						
Teacher survey						
Demographic and educational biographical information	1	Gender, age, highest educational attainment, professional degree	X	X	—	—
Professional education and attitudes toward the teaching profession	3	Teaching experience, university major	X	—	—	—
School structure and organizational characteristics	3	School atmosphere, communication with fellow teachers, willingness to innovate among the teaching staff	X	—	—	—
Structure, process, and organizational characteristics of subject teaching	5	Classroom disturbances, constructivist beliefs, cognitive activation	X	—	—	—
Spelling instruction	3	Spelling lessons (time to practice, materials, and methods)	X	—	—	—
Use of phonetic tables in teaching	1	Use of phonics tables (time, form, frequency, and own opinion)	X	X	—	—
Collaborative teaching	1	Students with and without special educational needs (preparation, change of teaching)	X	—	—	—
Resources	3	Time for reading instruction/reading activities, computer use in reading instruction, classroom library, reading specialist(s) on staff	X	—	—	—
Design of reading instruction	3	Differentiation in reading lessons, text types in reading instruction, methods in reading instruction, teaching reading strategies, support for weak and strong readers	X	—	—	—
School principal survey						
Social environment of the school and school climate	1	School spirit among teachers, parents, and students	X	—	—	—
Structure and organizational characteristics of school and teaching	3	Number of teaching days per year/week, performance orientation of the school, classroom management: organization of learning groups	X	—	—	—
Resources and equipment	3	Free meals at the school, homework supervision, number of computers/e-books available, problems with the school's personnel and material equipment, all-day school	X	—	—	—
Additional support services offered by the school	3	School support and care services	X	X	—	—
Measures for the promotion of reading	3	Systematic development of reading skills/strategies, measures to promote reading	X	—	—	—
Inclusion	2	*Inklusionsschule*, e.g., schools that include students both with and without special needs (implementation, year, number of students, specialized support without diagnoses, type of support needs, persons involved, framework)	X	—	—	—
Immigration	2	First-generation immigrants and refugees (number of students, organization, challenges)	X	X	—	—
School profile and profile classes	3	School profile (type, goals), profile classes (type and goals)	X	—	—	—
Information about students	3	Number of students, social backgrounds, German as native language	X	X	—	—
Information about teachers	4	Problematic behavior by teachers, age composition, teacher cooperation	X	X	—	—
Information about school principal	2	Gender, experience, areas of activity of school administration	X	—	—	—

Region, combining rural, urban, and metropolitan regions; ATR, academic track recommendation; EoS, enjoyment of school; RI, reading interest

Notes. "X" indicates that at least one principal component was selected from the respective group for the above-named outcome; "—" indicates that none were selected. We followed the PIRLS 2016 codebook for the grouping of the variables. A table showing explicitly which principal component of a group of variables was selected is provided as Online Appendix: https://osf.io/dyj3h/?view_only=0de4096d6090400694b99a2d0cdfaeb7.[8]

learning and social behavior in class. In addition to grades, the first two principal components from the skills group were selected to explain both the region and academic track recommendation. Furthermore, the first principal component of parental support behavior (e.g., parents' interest in the child's activities) also could serve as a regional classifier and predictor of reading interest.

Additionally, school characteristics also were selected: The first two principal components of subject-specific instruction in the classroom were critical factors in describing the regions as well as enjoyment of school and reading interest. Variables that made up this principal component included classroom reading activities, perceived classroom management, and perceived teacher support. In addition, the first two principal components of use of computers and internet in the classroom were selected to explain the region and enjoyment of school.

Parent Survey Data. Fewer principal components were selected from the parent data overall, and only two (student characteristic) principal components were simultaneously selected as classifiers for region and predictors of outcome variables. The second principal component of social background explained the regions and academic track recommendations. In addition, the first principal component of the reference person influenced both the regions and enjoyment of school. No variable explained both regions and reading interest.

Interestingly, principal components from students' social background explaining both a region and one of the outcomes were only the case for parents and not for the student survey. Furthermore, although cultural background was selected to explain the regions, it was not selected to explain the outcomes.

Selected School-Level Information. Teacher Survey Data. For teachers, all principal components were selected to explain the regions, but only very few were selected for academic track recommendation, and none were selected for enjoyment of school and reading interest. The first principal component of demographic and educational biographical information as well as the use of phonics tables in teaching helped to classify the regions and predicted variation in academic track recommendations.

Principal Survey Data. As with the teacher survey data, all principal components were selected to explain the regions, but only a few also explained academic track recommendations while not explaining enjoyment of school or reading interest. The first principal components of additional support services offered by the school, immigration, information about students, and information about teachers were selected to explain the regions and the academic track recommendation.

These selection results were quite surprising. Although LASSO is designed to handle many variables, our results raised the question of whether LASSO can handle hierarchical data well.

Post Hoc Regression Results Categorizing Variables into Directly Influenced by Policy, Indirectly Influenced by Policy, and Not Influenced by Policy. To further highlight whether the regional effects measured in Table 3 can be explained by variables which are affected by policy decisions, we categorized the principal components into those that can be a) directly influenced by policy, b) indirectly influenced by policy, and c) not at all influenced by policy. Each of the four authors first categorized the variables separately. For the 40 variable groups, in 11 cases (28%) all the authors agreed on the categorization. In 21 cases (52%), at least three of the authors agreed on the categorization. In only 20% of the cases did only two of the authors agree on the categorization. The categorization can be seen in the Online Appendix: https://osf.io/dyj3h/?view_only=0de4096d6090400694b99a2d0cdfaeb7. Then, instead of separating the data into student and school levels, we separated them into the three categories and redid the selection of principal components. The results show that classifying the principal components according to their potential to be influenced by policy helped to clarify regional differences in academic track recommendations, enjoyment of school, and interest in reading. Table 5 presents the OLS and post-LASSO regression results. The baseline model (model 1) is included for ease of reference.

This extends our understanding as follows: For academic track recommendations, nonpolicy variables most significantly reduced regional coefficients from $\beta_{rural}=.109$ to .018 and $\beta_{metropolitan}=.151$ to .050 in model 9. Direct policy variables also reduced these coefficients, but to a lesser extent, whereas indirect policy variables left the metropolitan coefficients significant in models 12 and 13, suggesting an inadequate explanation of regional differences. For students' subjective well-being, direct policy variables were most effective in reducing the coefficients in models 10 and 11, but the inclusion of indirect policy variables also significantly reduced the regional coefficients, especially in model 13. In the case of reading interest, nonpolicy variables initially reduced regional differences substantially, but indirect policy variables suggested a potential positive impact on reading interest in metropolitan areas (model 12), although their overall impact varied by region. Thus, although policy-related variables could explain some regional differences, nonpolicy variables played a significant role, particularly for academic track recommendations. Direct policy reforms could reduce regional differences among students in primary school, but student-targeted learning interventions might be more effective at increasing students' academic success.

TABLE 5
Ordinary-least-squares and post-LASSO regression results

Factor	OLS baseline	Post-LASSO: Not influenced by policy		Post-LASSO: Direct		Post-LASSO: Indirect	
	Model 1	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13
Panel A	Dependent variable: *academic track recommendation*						
Rural region	0.109*	0.027	0.018	0.051	0.063	0.071	0.046
	(0.045)	(0.041)	(0.041)	(0.055)	(0.055)	(0.043)	(0.041)
Metropolitan region	0.151**	0.054	0.050	0.084	0.080	0.095*	0.086*
	(0.053)	(0.042)	(0.040)	(0.059)	(0.058)	(0.044)	(0.039)
c_i^r	No	Yes	Yes	Yes	Yes	Yes	Yes
c_i^{ry}	No	No	Yes	No	Yes	No	Yes
Avg. observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608
R^2	0.093	0.479	0.500	0.229	0.296	0.197	0.364
R^2_{adj}	0.089	0.473	0.494	0.216	0.284	0.192	0.359
BIC (null)	−204.5	−1,978.127	−2,083.712	−425.178	−730.089	−594.488	−1,380.788
Variables	18	40	45	61	63	23	27
Panel B	Dependent variable: *enjoyment of school*						
Rural region	0.302**	0.107	0.109	0.012	0.016	0.011	0.087
	(0.107)	(0.102)	(0.100)	(0.133)	(0.103)	(0.114)	(0.082)
Metropolitan region	0.239***	0.075	0.087	0.058	0.053	0.022	0.050
	(0.131)	(0.102)	(0.100)	(0.133)	(0.103)	(0.121)	(0.094)
c_i^r	No	Yes	Yes	Yes	Yes	Yes	Yes
c_i^{ry}	No	No	Yes	No	Yes	No	Yes
Avg. observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608
R^2	0.035	0.236	0.259	0.170	0.372	0.193	0.284
R^2_{adj}	0.030	0.228	0.250	0.156	0.361	0.188	0.280
BIC (null)	13.480	−632.643	−704.727	−165.553	−1125.944	−575.488	−992.106
Variables	18	40	44	61	64	23	24
Panel C	Dependent variable: *reading interest*						
Rural region	0.210***	−0.030	−0.044	−0.084	−0.085	0.081	−0.028
	(0.109)	(0.091)	(0.080)	(0.131)	(0.143)	(0.097)	(0.067)
Metropolitan region	0.253	0.032	0.041	0.045	0.041	0.132	0.027
	(0.138)	(0.096)	(0.105)	(0.140)	(0.105)	(0.112)	(0.076)
c_i^r	No	Yes	Yes	Yes	Yes	Yes	Yes
c_i^{ry}	No	No	Yes	No	Yes	No	Yes
Avg. observations	3,608	3,608	3,608	3,608	3,608	3,608	3,608
R^2	0.024	0.267	0.378	0.164	0.324	0.104	0.441
R^2_{adj}	0.019	0.259	0.371	0.150	0.312	0.099	0.437
BIC (null)	53.948	−774.885	−1,339.976	−141.519	−866.594	−207.617	−1,861.773
Variables	18	40	42	61	64	23	24

OLS, ordinary least squares; LASSO, least absolute shrinkage and selection operator; BIC, Bayesian information criterion

Notes. The table shows the OLS and post-LASSO regression results for the three outcome variables: *academic track recommendation* (panel A), *enjoyment of school* (panel B), and *reading interest* (panel C) using the school bootstrap sample. That is, we resampled the schools before imputing missing data, resulting in, on average, 3,608 students (Avg. observations) across the imputed datasets. First, model 1 presents the OLS (baseline) regression results including only the state fixed effects as additional variables. Next, the table shows post-LASSO regressions at the student level, in which we have included selected principal components for regions (model 2) and for regions and the respective outcomes (nodel 3). The same is done for school-level principal components (models 4 and 5) and selected student- and school-level principal components (all-in-one, models 6 and 7). Full regression results with all coefficients can be seen in the online Appendix. Regression results without state fixed effects can be found in the Online Appendix: https://osf.io/dyj3h/?view_only=0de4096d6090400694b99a2d0cdfaeb7.

*$p < .05$; **$p < .01$; ***$p < .1$.

Discussion

This study examined regional differences among a representative sample of fourth grade elementary school children in Germany. To do this, we focused on the academic track recommendation because it is a measure of students' current academic performance and has long-term consequences for their subsequent school and career trajectories (Biewen & Tapalaga, 2017). Additionally, to gain broader insight into students' overall educational experience, we further looked at students' enjoyment of school as well as their general reading interest.

We followed the literature by dividing regions into three groups: metropolitan, urban, and rural areas. Overall, in RQ1, we found that regional differences in school performance can be observed not only in countries with above-average income inequality, as measured by the Gini coefficient (such as the United States and Israel), but also in Germany (which exhibits lower income inequality and therefore a below-average Gini coefficient). Using urban regions as a comparison group, students from metropolitan and rural areas were more likely to receive an academic track recommendation. Additionally, we found that students from metropolitan areas experienced greater school enjoyment, whereas students from rural areas had higher reading interest. Although regions are influenced by unique social and economic constructs and should not be expected to be homogeneous, we found some general regional effects in Germany. Our results underline that the composition of regional areas—consisting of residents, schools, teachers, and school principals—is inherently shaped by the underlying socioeconomic constructs of society. In this context, the inclusion of state fixed effects was crucial because it allowed us to disentangle regional influences from broader state-level heterogeneity in education policy and resource allocation.

Our results on the elementary school graduation variable *academic track recommendation* contrasted with the achievement results of Miller and Votruba-Drzal (2013) and Roscigno et al. (2006). They found lower academic achievement among students from metropolitan and rural areas in kindergarten and secondary school. This difference may be rooted in country-specific characteristics of Germany and the United States. For example, German regions with between 15,000 and 100,000 inhabitants (reference category) include many cities, not just suburban areas, like the United States. Additionally, the difference also could be rooted in the fact that we focused on students' academic track recommendation and not on an achievement test score. For students' school-related well-being, our results echoed the results of Gross-Manos and Shimoni (2020). Additionally, the positive effect of the rural region is in line with González-Carrasco et al. (2019) and Rees et al. (2017). Unfortunately, these studies did not differentiate between urban and metropolitan areas, which is why we cannot compare coefficients for the metropolitan region directly. Lastly, for reading interest, our results yielded higher reading interest among rural students compared with urban students in the baseline results. At the same time, there was a tendency for lower reading interest once we included additional variables in the regressions. This negative effect would align with research arguing that children in rural areas are less likely to read (Cantrell et al., 2018).

Once we included the additional information from the data from students, parents, teachers, and school principals using either the post-selection regression (RQ2a) or the post-double-selection regressions (RQ2b), we found that student-level information overwrote the regional effects. This suggests that the observed regional differences are primarily due to the characteristics observed in PRILS 2016 that describe the regions. Additionally, the LASSO included all principal components at the school level explaining the regions, leading to a very high number of principal components in the school-level and all-in-one post-LASSO regressions, increased standard errors, and a worse Bayesian information criterion than the student-level information alone. Thus, having too much level 2 information in hierarchical data might make it more difficult for LASSO to select incremental information.

The selected variables that helped explain the regional effects of the rural and metropolitan regions compared with the urban regions on the student and school levels can be attributed to student and school characteristics (RQ3). Regional differences, as well as students' academic achievement, are related to student characteristics such as social background, skills, and learning and social behaviors at school (student level) as well as the composition of the student body (information about students) and immigration (school level). However, the results also identified some school characteristics for the regions and the outcomes that (e.g.) policymakers can influence: promoting increased parental support behavior (students: parental support behavior; parents: elementary school years) to increase academic track recommendations and reading interest, setting helpful subject-specific teaching standards in the classroom to improve students' enjoyment of school and reading interest, and teacher training on good use of computers and the internet in the classroom to increase students' enjoyment of school (all student level). Again, from the school level, these are creating incentives for good teachers (demographic and educational biographical information, information about teachers) to work in disadvantaged regions as well as (financially) enabling additional support services offered by the school.

Our results are in line with the educational policy literature showing how the unequal distribution of (highly skilled) teachers and resources affects student performance (Bertoni et al., 2023; Edwards, 2021; Fuller & Liang, 1996; Gershenson et al., 2017; Lankford et al., 2002; Loeb et al., 2010; Reininger, 2012). Our post hoc regression underscores this point by showing that the inclusion of directly and indirectly policy-related principal components reduces regional differences in a similar way to nonpolicy principal components. The indirect policy variables are an exception when it comes to explaining academic track recommendations, for which the nonpolicy variables reduce regional differences the most. Thus, we have highlighted that policy choices matter for regional differences in academic achievement, students' school well-being, and reading interest and have explored which characteristics researchers and policymakers should focus on. This is of high importance because small (dis)advantages early in life can lead to gross inequalities in adulthood (DiPrete & Eirich, 2006; Elder et al., 2003).

Limitations and Strengths

When interpreting the results, certain aspects of this study warrant attention. Our measure of academic achievement was a proxy and not subject specific or a specific test. Moreover, PIRLS 2016 includes only one affective component of well-being. Therefore, we could not examine multiple components of well-being. Additionally, because PIRLS 2016 does not contain postal codes for data-protection reasons, we were restricted to the collected variables and could not include economic data at the regional, district, or school level. Our taxonomy of rural, (sub)urban, and metropolitan regions based on the number of inhabitants is a consequence of these limitations. Lastly, the question of "Which policy is most effective?" or "How much would implementing policy XYZ alter academic achievement, student well-being, and reading interest?" is not within the scope of this study.

Despite these limitations, the large sample size of the PIRLS 2016 data, their representativeness, and standardized data collection are clear strengths for providing meaningful insights on regional differences in academic achievement, well-being, and (reading) interest among elementary school children. Additionally, we systematically expanded the level of knowledge by (a) focusing on elementary schools in Germany and (b) focusing on different central outcomes. Additionally, we went beyond just describing whether regional differences exist and examined whether school-related data can fully explain where these differences come from. Furthermore, we applied cutting-edge machine learning methods that simultaneously allowed for the theoretical inclusion of a rich set of control variables while also avoiding overfitting when selecting the essential variables. Applying selection methods for both the outcomes and the main explanatory variables further decreased the probability of omitted variable bias in the found effects of the explanatory variables. Lastly, we used a cutting-edge method to focus on more than just a few variables and instead include the richness of the entire PIRLS 2016 dataset to explain regional differences in several outcome variables. Thereby, we illustrate how a combination of several methods can be used to analyze a large dataset in its entirety. In the future, if our procedure is followed, it will be possible to determine, when analyzing large data, whether the results are robust not only to a subjectively chosen set of a few control variables but also to the entire set of data collected.

We showed that when analyzing student data across different regions, it is essential to include teacher and school principal data to control for regional differences in country-wide analyses of academic achievement. Merely examining student and parent data is insufficient. In addition, our analysis highlights that regional differences also exist in a country with a below-average Gini coefficient and that more attention is needed to equalize regional differences beyond mitigating income inequality. A lot of the selected school characteristics are directly influenced by policy decisions. Because our study does not claim a causal relationship, further (quasi-) experimental research is needed to take a closer look at the selected variables to identify specific variables that policymakers could optimize to improve academic achievement, students' school well-being, and (reading) interest.

Declaration of Use of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used GPT-4 and DeepL for language-focused editing of human-written text. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of this publication.

Funding

This work was supported by "From Prediction to Agile Interventions in the Social Sciences (FAIR, PROFILNRW-2020-068)." The sole responsibility for the content of this publication lies with the authors.

ORCID iD

Jakob Schwerter https://orcid.org/0000-0001-5818-2431

Notes

1. In our setting, German laws exacerbate this aspect. In order to plan and manage school capacities, the German federal states have implemented laws allowing families to choose an elementary school only within the school district in which the parents live. Parents have one to several schools to choose from within the district depending on its size. Only when important arguments are brought forward in a formal legal proceeding (sometimes requiring parents to hire a lawyer) can the school authorities grant an exemption from this rule (Mayer & Koinzer, 2020).

2. The Gini coefficient is a commonly used measure of income inequality. It ranges from 0 to 1, with 0 representing perfect equality (where everyone has the same income) and 1 representing perfect inequality (where one person has all the income). The Organisation for Economic Co-operation and Development (OECD) collects data on income inequality in its member countries.

3. The exchange rate of 1 euro to the U.S. dollar has fluctuated since 2001 from 0.8956 (minimum) to 1.4708 (maximum) to 1.0530 in 2022.

4. Obviously, teacher or family information does not "predict" a region in any real sense. We simply mean whether we could accurately predict the region in which a student lived based on family information, that is, important components of the respective regional profiles.

5. In Germany, an official typology of municipalities based on the number of inhabitants exists (see the taxonomy of the Bundesinstitut für Bau-, Stadt-, und Raumforschung (BBSR; https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/gemeinden/StadtGemeindetyp/StadtGemeindetyp.html). A town (*Kleinstädt*) is defined as having 5,000–20,000 inhabitants, which can be subcategorized into small *Kleinstädte*, <10,000 inhabitants, and greater *Kleinstädte*, >10,000 inhabitants. In our tripartition of regions, we set the cutoff at 15,000 because we would consider the majority of *Kleinstädte* as situated in rural regions. Above 15,000 inhabitants, we would speak of a mostly urban environment. The BBSR categorizes towns with 20,000–100,000 inhabitants as midsized cities (*Mittelstädte*). We follow the BBSR's definition of large cities (*Großstadt*), which is defined as >100,000 inhabitants, with our definition of a metropolitan region. In contrast, the U.S. National Center for Education Statistics (NCES) uses an urban-centered locale system. This includes urban areas, both large (population of ≥250,000) and medium (population of 50,000–249,999); suburban locales located on the outskirts of these large and medium-sized cities; towns, which may be fringe (close to urban centers), distant, or remote; and rural categories that mirror the town designations in terms of proximity to urban centers (see the taxonomy of the U.S. NCES https://nces.ed.gov/surveys/annualreports/topical-studies/locale/definitions).

6. While in the United States only schools with a high-poverty student body have free meals available to all students (due to a federal funding stream), free meals in German schools are not provided just in high-poverty regions but are more of a perk of high-resource schools.

7. Regression results without state fixed effects can be found in the Online Appendix: https://osf.io/dyj3h/?view_only=0de4096d6090400694b99a2d0cdfaeb7.

8. It is not straightforward in our setting to test whether the regional effects have changed significantly across equations, that is, with the inclusion of the selected principal components. Unfortunately, we were not able to find any standard cross-equation testing procedure in the literature that captures the peculiarities of our application—namely a cross-equation test for multiple imputed, bootstrapped post-selection regression estimates for nested models. Therefore, we set out our test strategy in the Online Appendix: https://osf.io/dyj3h/?view_only=0de4096d6090400694b99a2d0cdfaeb7. We assess that the change of the regional coefficients is particularly significant when including the selected principal components from the student-level data.

Note: This manuscript was accepted under the editorship of Dr. Kara Finnigan.

References

Abramitzky, R., Lavy, V., & Pérez, S. (2021). The long-term spillover effects of changes in the return to schooling. *Journal of Public Economics*, *196*, 104369. https://doi.org/10.1016/j.jpubeco.2021.104369

Akande, O., Li, F., & Reiter, J. (2017). An empirical comparison of multiple imputation methods for categorical sata. *American Statistician*, *71*(2), 162–170. https://doi.org/10.1080/00031305.2016.1277158

Aldrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit, and probit models* (Issue 45). Sage.

Alejo, R., & Piquer-Píriz, A. (2016). Urban vs. rural CLIL: An analysis of input-related variables, motivation and language attainment. *Language, Culture and Curriculum*, *29*(3), 245–262. https://doi.org/10.1080/07908318.2016.1154068

Asadi, I. A. (2018). Reading comprehension subgroups in arabic: A simple but not a multiplicative model. *Reading and Writing Quarterly*, *34*(4), 281–290. https://doi.org/10.1080/10573569.2017.1387835

Bacolod, M. (2007). Who teaches and where they choose to teach: College graduates of the 1990s. *Educational Evaluation and Policy Analysis*, *29*(3), 155–168. https://doi.org/10.3102/0162373707305586

Barrett, N., Carlson, D., Harris, D. N., & Lincove, J. A. (2022). When the walls come down: Evidence on charter schools' ability to keep their best teachers without unions and certification rules. *Educational Evaluation and Policy Analysis*, *44*(2), 283–312. https://doi.org/10.3102/01623737211047265

Bastian, K. C., & Fortner, C. K. (2020). Is less more? Subject-area specialization and outcomes in elementary schools. *Education Finance and Policy*, *15*(2), 357–382. https://doi.org/10.1162/edfp_a_00278

Beck, N. (2019). Estimating grouped data models with a binary-dependent variable and fixed effects via a logit versus a linear probability model: The impact of dropped units. *Political Analysis*, *22*(18), 139–145. https://doi.org/10.1017/pan.2019.20

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, *81*(2), 608–650. https://doi.org/10.1093/restud/rdt044

Bertoni, E., Elacqua, G., Hincapié, D., Méndez, C., & Paredes, D. (2023). Teachers' preferences for proximity and the implications for staffing schools: Evidence from peru. *Education Finance and Policy*, *18*(2), 181–212. https://doi.org/10.1162/edfp_a_00347

Biewen, M., & Tapalaga, M. (2016). *Life-cycle educational choices: Evidence for two German cohorts* (IZA Discussion Papers No. 9699). Institute of Labor Economics.

Biewen, M., & Tapalaga, M. (2017). Life-cycle educational choices in a system with early tracking and "second chance" options. *Economics of Education Review*, *56*, 80–94. https://doi.org/10.1016/j.econedurev.2016.11.008

Bornstein, M. H., Putnick, D. L., Heslington, M., Gini, M., Suwalsky, J. T. D., Venuti, P., de Falco, S., Giusti, Z., & de Galperín, C. Z. (2008). Mother–child emotional availability in ecological perspective: Three countries, two regions, two genders. *Developmental Psychology*, *44*(3), 666–680. https://doi.org/10.1037/0012-1649.44.3.666

Burchinal, M., Vernon-Feagans, L., Vitiello, V., Greenberg, M., Vernon-Feagans, L., Greenberg, M., Cox, M., Blair, C., Burchinal, M., Willoughby, M., Garrett-Peters, P., Mills-Koonce, R., & Ittig, M. (2014). Thresholds in the association between child care quality and child outcomes in rural preschool children. *Early Childhood Research Quarterly*, *29*(1), 41–51. https://doi.org/10.1016/j.ecresq.2013.09.004

Byrnes, J. P., & Miller, D. C. (2007). The relative importance of predictors of math and science achievement: An opportunity-propensity analysis. *Contemporary Educational Psychology*, *32*(4), 599–629. https://doi.org/10.1016/j.cedpsych.2006.09.002

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S. L., Hiebert, J., & Bakker, A. (2020). Maximizing the quality of learning opportunities for every student. *Journal for Research in Mathematics Education*, *51*(1), 12–25. https://doi.org/10.5951/jresematheduc.2019.0005

Cantrell, S. C., & Rintamaa, M. (2020). The nature of rural high school students' reading engagement. *Reading and Writing Quarterly*, *36*(4), 297–319. https://doi.org/10.1080/10573569.2019.1635543

Cantrell, S. C., Rintamaa, M., Anderman, E. M., & Anderman, L. H. (2018). Rural adolescents' reading motivation, achievement and behavior across transition to high school. *Journal of Educational Research*, *111*(4), 417–428. https://doi.org/10.1080/00220671.2017.1284737

Caudill, S. B. (1988). Practitioners corner: An advantage of the linear probability model over probit or logit. *Oxford Bulletin of Economics and Statistics*, *50*(4), 425–427. https://doi.org/10.1111/j.1468-0084.1988.mp50004005.x

Chatla, S., & Shmueli, G. (2013). Linear probability models (LPM) and big data: The good, the bad, and the ugly. *SSRN Electronic Journal*, 1–45. https://doi.org/10.2139/ssrn.2353841

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, *106*(4), 855–902. https://doi.org/10.1257/aer.20150572

Cooper, B. R., Moore, J. E., Powers, C. J., Cleveland, M., & Greenberg, M. T. (2014). Patterns of early reading and social skills associated with academic success in elementary school. *Early Education and Development*, *25*(8), 1248–1264. https://doi.org/10.1080/10409289.2014.932236

Cutler, D. M., & Glaeser, E. L. (1997). Are ghettos good or bad? *Quarterly Journal of Economics*, *112*(3), 826–872. https://doi.org/10.1162/003355397555361

Diener, E., & Ryan, K. (2009). Subjective well-being: A general overview. *South African Journal of Psychology*, *39*(4), 391–406. https://doi.org/10.1177/008124630903900402

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*, *32*, 271–297. https://doi.org/10.1146/annurev.soc.32.061604.123127

Eccles, J. S., & Roeser, R. (2009). Schools, academic motivation, and stage-environment fit. In R. M. Lerner, & L. Steinberg (Eds.), *Handbook of adolescent psychology: Individual bases of adolescent development* (1st ed., pp. 404–434). Wiley.

Education Reporting Consortium. (2006). *Bildung in Deutschland. Ein indikatoren- gestützter Bericht mit einer Analyse zu Bildung und Migration* [Education in Germany: An indicator-based report with an analysis of education and migration]. Konsortium Bildungsberichterstattung.

Edwards, D. S. (2021). Just out of reach? Unrestrained supply, constrained demand, and access to effective schools in and around detroit. *Educational Evaluation and Policy Analysis*, *43*(3), 391–418. https://doi.org/10.3102/0162373721996738

Elder, G., Johnson, M., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer, & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 3–19). Springer. https://doi.org/10.1007/978-0-306-48247-2_1

European Commission. (2010). *Communication from the commission Europe 2020: A strategy for smart, sustainable and inclusive growth*. http://ec.europa.eu/news/economy/100303_en.htm

Farrell, C. R. (2008). Bifurcation, fragmentation or integration? The racial and geographical structure of US metropolitan segregation, 1990–2000. *Urban Studies*, *45*(3), 467–499. https://doi.org/10.1177/0042098007087332

Fiorilli, C., De Stasio, S., Di Chiacchio, C., Pepe, A., & Salmela-Aro, K. (2017). School burnout, depressive symptoms and engagement: Their combined effect on student achievement. *International Journal of Educational Research*, *84*, 1–12. https://doi.org/10.1016/j.ijer.2017.04.001

Foster, T. B. (2017). Decomposing American immobility: Compositional and rate components of interstate, intrastate, and intracounty migration and mobility decline. *Demographic Research*, *37*(1), 1515–1548. https://doi.org/10.4054/DemRes.2017.37.47

Fuller, B., & Liang, X. (1996). Market failure? Estimating inequality in preschool availability. *Educational Evaluation and Policy Analysis*, *18*(1), 31–49. https://doi.org/10.3102/01623737018001031

García-Crespo, F. J., Fernández-Alonso, R., & Muñiz, J. (2021). Academic resilience in European countries: The role of teachers, families, and student profiles. *PLoS One*, *16*, 1–20. https://doi.org/10.1371/journal.pone.0253409

German Federal Statistical Office. (2010). Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund—Ergebnisse des Mikrozensus 2009 [Population and employment. Population with a migration background—results of the 2009 microcensus]. In *Fachserie 1, Reihe 2.2, Statistisches Bundesamt*. https://leipzig.verband-binationaler.de/fileadmin/user_upload/_imported/fileadmin/Dokumente/PDFs/Mikrozensus_2013.pdf

German Federal Statistical Office. (2021). *Expenditure per pupil in puplic general and vocational schools*. Educational Finance, Promotion of Education and Training. https://leipzig.verband-binationaler.de/fileadmin/user_upload/_imported/fileadmin/Dokumente/PDFs/Mikrozensus_2013.pdf

Gershenson, S., Jacknowitz, A., & Brannegan, A. (2017). Are student absences worth the worry in U.S. primary schools? *Education Finance and Policy*, *12*(2), 137–165. https://doi.org/10.1162/EDFP_a_00207

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1), 3–42. https://doi.org/10.1007/s10994-006-6226-1

Goldfarb, A., & Tucker, C. (2011). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, *30*(3), 389–404. https://doi.org/10.1287/mksc.1100.0583

Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, *150*(4), 700–709. https://doi.org/10.1037/xge0000920

González-Carrasco, M., Casas, F., Viñas, F., Malo, S., & Crous, G. (2019). The interplay between school and home location and its relationship with children's subjective well-being. *Children's*

Geographies, *17*(6), 676–690. https://doi.org/10.1080/14733285.2019.1635993
Govorova, E., Benítez, I., & Muñiz, J. (2020). How schools affect student well-being: A cross-cultural approach in 35 OECD countries. *Frontiers in Psychology*, *11*, 1–14. https://doi.org/10.3389/fpsyg.2020.00431
Gross-Manos, D., & Shimoni, E. (2020). Where you live matters: Correlation of child subjective well-being to rural, urban, and peripheral living. *Journal of Rural Studies*, *76*, 120–130. https://doi.org/10.1016/j.jrurstud.2020.04.009
Hanson, M. (2022, June 15). *U.S. public education spending statistics*. EducationData.org. https://educationdata.org/public-education-spending-statistics
Hanushek, E. A., Light, J. D., Peterson, P. E., Talpey, L. M., & Woessmann, L. (2022). Long-run trends in the U.S. achievement gap. *Education Finance and Policy*, *14*(46), 1–33. https://doi.org/10.1162/edfp_a_00383
Hascher, T. (2008). Quantitative and qualitative research approaches to assess student well-being. *International Journal of Educational Research*, *47*(2), 84–96. https://doi.org/10.1016/j.ijer.2007.11.016
Hastie, T., Tibshirani, R., & Friedman, J. (2009). Linear methods for regression. In T. Hastie, R. Tibshirani, & J. Friedman (Eds.), *The elements of statistical learning* (pp. 43–94). Springer. https://doi.org/10.1007/b94608
Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality and Quantity*, *43*(1), 59–74. https://doi.org/10.1007/s11135-007-9077-3
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
Horrace, W. C., & Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, *90*(3), 321–327. https://doi.org/10.1016/j.econlet.2005.08.024
Hübner, N., Wille, E., Cambria, J., Oschatz, K., Nagengast, B., & Trautwein, U. (2017). Maximizing gender equality by minimizing course choice options? Effects of obligatory coursework in math on gender differences in STEM. *Journal of Educational Psychology*, *109*(7), 993–1009. https://doi.org/10.1037/edu0000183
Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., Lankes, E.-M., McElvany, N., Stubbe, T. C., & Valtin, R. (2017). *IGLU 2016*. *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster; New York: Waxmann.
Ivey, G., & Johnston, P. H. (2013). Engagement with young adult literature: Outcomes and processes. *Reading Research Quarterly*, *48*(3), 255–275. https://doi.org/10.1002/rrq.46
Kaas, L., & Manger, C. (2012). Ethnic discrimination in Germany's labour market: A field experiment. *German Economic Review*, *13*(1), 1–20. https://doi.org/10.1111/j.1468-0475.2011.00538.x
Kirsch, I., de Jong, J., Lafontaine, D., McQueen, J., Mendelovitz, J., & Monseur, C. (2002). *Reading for change: Performance and engagement across countries*. OECD. https://doi.org/10.1787/9789264099289-en
Kleinkorres, R., Stang-Rabrig, J., & McElvany, N. (2023). Comparing parental and school pressure in terms of their relations with students' well-being. *Learning and Individual Differences*, *104*(2), 102288. https://doi.org/10.1016/j.lindif.2023.102288
Köller, O., Hasselhorn, M., Hesse, F. W., Maaz, K., Schrader, J., Solga, H., Spieß, C. K., & Zimmer, K. (2021). *Das Bildungswesen in Deutschland* [The education system in Germany]. Verlag Julius Klinkhardt https://doi.org/10.36198/9783838547855
Kosse, F., Deckers, T., Pinger, P., Schildberg-Hörisch, H., & Falk, A. (2020). The formation of prosociality: Causal evidence on the role of social environment. *Journal of Political Economy*, *128*(2), 434–467. https://doi.org/10.1086/704386
Krabel, S., & Flöther, C. (2014). Here today, gone tomorrow? Regional labour mobility of German university graduates. *Regional Studies*, *48*(10), 1609–1627. https://doi.org/10.1080/00343404.2012.739282
Krieg, J. M., Theobald, R., & Goldhaber, D. (2015). A foot in the door: Exploring the role of student teaching assignments in teachers' initial job placements. *Educational Evaluation and Policy Analysis*, *38*(2), 364–388. https://doi.org/10.3102/0162373716630739
Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language Learning*, *62*(4), 997–1023. https://doi.org/10.1111/j.1467-9922.2012.00719.x
Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, *24*(1), 37–62. https://doi.org/10.3102/01623737024001037
Lavy, V. (2021). The long-term consequences of free school choice. *Journal of the European Economic Association*, *19*(3), 1734–1781. https://doi.org/10.1093/jeea/jvab001
Lee, D., & Hosanagar, K. (2021). How do product attributes and reviews moderate the impact of recommender systems through purchase stages? *Management Science*, *67*(1), 524–546. https://doi.org/10.1287/mnsc.2019.3546
Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule [Estimating causal effects: Reflections on methods of causal analysis using context effects in schools.]. *Kolner Zeitschrift fur Soziologie und Sozialpsychologie*, *64*(1), 123–153. https://doi.org/10.1007/s11577-012-0158-5
Leiss, D., Plath, J., & Schwippert, K. (2019). Language and mathematics—Key factors influencing the comprehension process in reality-based tasks. *Mathematical Thinking and Learning*, *21*(2), 131–153. https://doi.org/10.1080/10986065.2019.1570835
Loeb, S., Kalogrides, D., & Horng, E. L. (2010). Principal preferences and the uneven distribution of principals across schools. *Educational Evaluation and Policy Analysis*, *32*(2), 205–229. https://doi.org/10.3102/0162373710369833
Long, Q., & Johnson, B. A. (2015). Variable selection in the presence of missing data: Resampling and imputation. *Biostatistics*, *16*(3), 596–610. https://doi.org/10.1093/biostatistics/kxv003
Lumley, T. (2021). *Analysis of complex survey samples*. R-Package Survey Version 4.1-1. http://r-survey.r-forge.r-project.org/survey/
Lumley, T. (2010). *Complex surveys: A guide to analysis using R* (pp. 1–271). Wiley.

Maaz, K., Artelt, C., Brugger, P., Buchholz, S., Kühne, S., Leerhoff, H., Rauschenbach, T., Rockmann, U., Roßbach, H.-G., Schrader, J., & Seeber, S. (2020). *Bildung in Deutschland* [Education in Germany 2020]. wbv Publication. https://doi.org/10.3278/6001820gw

Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, *110*, 63–73. https://doi.org/10.1016/j.jclinepi.2019.02.016

Mathis, W. (2003). Equity and adequacy challenges in rural schools and communities. *Journal of Education Finance*, *29*(2), 119–136. http://eric.ed.gov/ERICWebPortal/recordDetail?accno=ED478057

Mayer, M. (2019). *missRanger: Fast imputation of missing values*. R Package Version 2.1.3. https://doi.org/10.1093/bioinformatics/btr597

Mayer, T., & Koinzer, T. (2020). Schulwahl und Bildungsungleichheit [School choice and educational inequality]. *Pädagogische Rundschau*, *73*(3), 265–280. https://doi.org/10.3726/pr032019.0025

McBreen, M., & Savage, R. (2021). The impact of motivational reading instruction on the reading achievement and motivation of students: A systematic review and meta-analysis. *Educational Psychology Review*, *33*(3), 1125–1163. https://doi.org/10.1007/s10648-020-09584-4

McCormick, M. P., O'Connor, E. E., Cappella, E., & McClowry, S. G. (2013). Teacher–child relationships and academic achievement: A multilevel propensity score model approach. *Journal of School Psychology*, *51*(5), 611–624. https://doi.org/10.1016/j.jsp.2013.05.001

Miller, P., & Votruba-Drzal, E. (2013). Early academic skills and childhood experiences across the urban-rural continuum. *Early Childhood Research Quarterly*, *28*(2), 234–248. https://doi.org/10.1016/j.ecresq.2012.12.005

Monk, D. H. (2007). Recruiting and retaining high-quality teachers in rural areas recruiting and retaining high-quality teachers in rural areas. *The Future of Children*, *17*(1), 155–174. www.futureofchildren.org

Murray, J. S. (2018). Multiple imputation: A review of practical and theoretical findings. *Statistical Science*, *33*(2), 142–159. https://doi.org/10.1214/18-STS644

Musterd, S. (2016). Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of Urban Affairs*, *27*(3), 331–348. https://doi.org/10.1111/j.0735-2166.2005.00239.x

Organisation for Economic Co-operation and Development (OECD). (2016). *OECD factbook 2015–2016: Economic, environmental and social Statistics*.

Organisation for Economic Co-operation and Development (OECD). (2017). *Education at a glance 2017: OECD indicators*. https://www.oecd.org/content/dam/oecd/en/publications/reports/2017/09/education-at-a-glance-2017_g1g7f6b6/eag-2017-en.pdf

Quillian, L., Heath, A., Pager, D., Midtbøen, A. H., Fleischmann, F., & Hexela, O. (2019). Do some countries discriminate more than others? Evidence from 97 field experiments of racial discrimination in hiring. *Sociological Science*, *6*, 467–496. https://doi.org/10.15195/V6.A18

Rees, G., Tonon, G., Mikkelsen, C., & Rodriguez de la Vega, L. (2017). Urban-rural variations in children's lives and subjective well-being: A comparative analysis of four countries. *Children and Youth Services Review*, *80*, 41–51. https://doi.org/10.1016/j.childyouth.2017.06.056

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis*, *34*(2), 127–145. https://doi.org/10.3102/0162373711420864

Renninger, K. A., & Hidi, S. E. (2020). To level the playing field, develop interest. *Policy Insights from the Behavioral and Brain Sciences*, *7*(1), 10–18. https://doi.org/10.1177/2372732219864705

Revelle, W. (2020). *psych: Procedures for psychological, psychometric, and personality research*. https://cran.r-project.org/package=psych

Roscigno, V. J., & Crowley, M. L. (2001). Rurality, institutional disadvantage, and achievement/attainment. *Rural Sociology*, *66*(2), 268–292. https://doi.org/10.1111/j.1549-0831.2001.tb00067.x

Roscigno, V. J., Tomaskovic-Devey, D., & Crowley, M. (2006). Education and the inequalities of place. *Social Forces*, *84*(4), 2121–2145. https://doi.org/10.1353/sof.2006.0108

Rothbart, M. W., Schwartz, A. E., & Gutierrez, E. (2023). Paying for free lunch: The impact of CEP universal free meals on revenues, spending, and student health. *Education Finance and Policy*, *18*(4), 708–737. https://doi.org/10.1162/edfp_a_00380

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.

Salemink, K., Strijker, D., & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. *Journal of Rural Studies*, *54*, 360–371. https://doi.org/10.1016/j.jrurstud.2015.09.001

Sánchez, A. C., & Andrews, D. (2011). Residential mobility and public policy. *OECD Journal: Economic Studies*, *2011*, 185–206. https://doi.org/10.1787/eco_studies-2011-5kg0vswqt240

Schmitt, S. A., Pratt, M. E., Korucu, I., Napoli, A. R., & Schmerold, K. L. (2018). Preschool classroom quality and social-emotional functioning: Findings across geographic regions. *Early Childhood Research Quarterly*, *43*, 11–22. https://doi.org/10.1016/j.ecresq.2017.11.002

Schwerter, J., Gurtskaia, K., Romero, A., Zeyer-Gliozzo, B., & Pauly, M. (2024). Evaluating tree-based imputation methods as an alternative to MICE PMM for drawing inference in empirical studies. *arXiv Preprint arXiv:2401.09602*.

Schwerter, J., Stang-Rabrig, J., Kleinkorres, R., Bleher, J., Doebler, P., & McElvany, N. (2024). Importance of students' social resource factors for academic achievement and well-being in elementary school. *European Journal of Psychology of Education*, *39*, 4515–4552. https://doi.org/10.1007/s10212-024-00877-8

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology. *American Psychologist*, *55*(1), 5–14. https://doi.org/10.1037/0003-066X.55.1.5

Seligman, M. E. P., Ernst, R. M., Gillham, J., Reivich, K., & Linkins, M. (2009). Positive education: Positive psychology and classroom interventions. *Oxford Review of Education*, *35*(3), 293–311. https://doi.org/10.1080/03054980902934563

Skinner, A. C., & Slifkin, R. T. (2007). Rural/urban differences in barriers to and burden of care for children with special health care needs. *Journal of Rural Health*, *23*(2), 150–157. https://doi.org/10.1111/j.1748-0361.2007.00082.x

Snow, C. E., & Sweet, A. P. (2003). Reading for comprehension. In A. P. Sweet, & E. Snow (Eds.), *Rethinking reading comprehension*. Guilford Press.

Steinmayr, R., Heyder, A., Naumburg, C., Michels, J., & Wirthwein, L. (2018). School-related and individual predictors of subjective well-being and academic achievement. *Frontiers in Psychology*, *9*, 1–16. https://doi.org/10.3389/fpsyg.2018.02631

Stekhoven, D. J., & Bühlmann, P. (2012). Missforest-Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Strayhorn, T. L. (2009). Different folks, different hopes. *Urban Education*, *44*(6), 710–731. https://doi.org/10.1177/0042085908322705

Thomä, J. (2023). An urban-rural divide (or not?): Small firm location and the use of digital technologies. *Journal of Rural Studies*, *97*, 214–223. https://doi.org/10.1016/j.jrurstud.2022.12.020

Tymms, P., Merrell, C., & Bailey, K. (2018). The long-term impact of effective teaching. *School Effectiveness and School Improvement*, *29*(2), 242–261. https://doi.org/10.1080/09243453.2017.1404478

UN Educational, Scientific and Cultural Organization (UNESCO). (2005). Education for all. In *Global monitoring report: Literacy for life*. https://doi.org/10.4324/9781315688213

UN Statistics Division. (2017). *Population density and urbanization*. https://unstats.un.org/unsd/demographic/sconcerns/densurb/densurbmethods.htm#B

Vaknin-Nusbaum, V., Nevo, E., Brande, S., & Gambrell, L. (2018). Developmental aspects of reading motivation and reading achievement among second grade low achievers and typical readers. *Journal of Research in Reading*, *41*(3), 438–454. https://doi.org/10.1111/1467-9817.12117

van Buuren, S. (2018). *Flexible imputation of missing data*. CRC Press. https://stefvanbuuren.name/fimd/

Vu, T., Van Magis-Weinberg, L., Jansen, B. R. J., van Atteveldt, N., Janssen, T. W. P., Lee, N. C., van der Maas, H. L. J., Raijmakers, M. E. J., Sachisthal, M. S. M., & Meeter, M. (2022). Motivation-achievement cycles in learning: A literature review and research agenda. *Educational Psychology Review*, *34*(1), 39–71. https://doi.org/10.1007/s10648-021-09616-7

Walgermo, B. R., Foldnes, N., Uppstad, P. H., & Solheim, O. J. (2018). Developmental dynamics of early reading skill, literacy interest and readers' self-concept within the first year of formal schooling. *Reading and Writing*, *31*(6), 1379–1399. https://doi.org/10.1007/s11145-018-9843-8

Wigfield, A., & Eccles, J. S. (2020). 35 years of research on students' subjective task values and motivation: A look back and a look forward. In A. J. Elliot (Ed.), *Advances in motivation science* (*Vol. 7*, pp. 161–198). Elsevier. https://doi.org/10.1016/bs.adms.2019.05.002

Wigfield, A., Gladstone, J. R., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives*, *10*(3), 190–195. https://doi.org/10.1111/cdep.12184

Wigfield, A., Guthrie, J. T., Tonks, S., & Perencevich, K. C. (2004). Children's motivation for reading: Domain specificity and instructional influences. *Journal of Educational Research*, *97*(6), 299–310. https://doi.org/10.3200/JOER.97.6.299-310

Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, *45*(5), 432–445. https://doi.org/10.1002/pits

Authors

JAKOB SCHWERTER is a postdoctoral researcher at the Center for Research on Education and School Development (IFS) and at the Chair of Statistical Methods in Social Sciences, both at TU Dortmund University. His research lies at the intersection of psychology, education, statistics, data science, and economics.

JOHANNES BLEHER is a postdoctoral researcher at the University of Hohenheim. His work focuses on econometrics, data science, and machine learning methods, with a particular emphasis on asset pricing.

PHILIPP DOEBLER is a full professor of statistical methods in social sciences at the Department of Statistics, TU Dortmund University. His recent work focuses on machine learning in measurement, meta-analysis, and educational data science.

NELE MCELVANY is executive director of the Center for Research on Education and School Development (IFS) and vice president for research at TU Dortmund University. As a professor of empirical educational research, she focuses on educational processes from psychological and pedagogic perspectives. She is a psychologist by training, and her research addresses individual, social, and institutional conditions of educational processes and outcomes.