

Lange, Simon; von Werder, Marten

**Conference Paper**

## Tracking and the Intergenerational Transmission of Education: Evidence from a Natural Experiment

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Economics of Education II, No. F14-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Lange, Simon; von Werder, Marten (2016) : Tracking and the Intergenerational Transmission of Education: Evidence from a Natural Experiment, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Economics of Education II, No. F14-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145784>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

# Tracking and the Intergenerational Transmission of Education: Evidence from a Natural Experiment

29 February, 2016

## **Abstract**

Proponents of tracking argue that the creation of more homogeneous classes increases efficiency while opponents fear that tracking aggravates initial differences between students. We estimate the effects on the intergenerational transmission of education of a reform that delayed tracking by two years in one of Germany's federal states. While the reform had no effect on educational outcomes *on average*, it increased educational attainment among individuals with uneducated parents and decreased attainment among individuals with educated parents. The reform thus lowered the gradient between parental education and own education. The effect is driven entirely by changes in the gradient for males.

**Keywords:** tracking; educational institutions; educational inequality; equality of opportunity; intergenerational mobility.

**JEL Classification Numbers:** I21, I24, I28, J62.

# 1 Introduction

Can education policies help level the playing field between students from different social backgrounds? And if so, how? One major design feature of education systems that has frequently been related to equality of opportunity is tracking, the practice of grouping students by ability (Betts, 2010). Countries differ widely in the way they track students<sup>1</sup> but almost all education systems in practice feature some kind of tracking. Proponents argue that the creation of more homogeneous classes increases efficiency by allowing educators to tailor lessons to students' specific needs. Opponents, on the other hand, fear that misclassification of students is often rife—especially when students are tracked at an early age—and that tracking aggravates initial differences.

In this paper we exploit a policy reform implemented during the 1970s in one of Germany's federal states, Lower Saxony. While most German states continued the tracking students after fourth grade, the reform shifted the timing of tracking from grade four to grade six (roughly ages ten and twelve, respectively). This was achieved through the introduction between 1972 and 1982 of a new intermediate school, the *orientation stage*<sup>2</sup> (henceforth, OS). We investigate the effects of this reform on (1) levels of educational attainment and (2) the intergenerational transmission of education, what we will refer to as the *gradient* between parental education and own education.<sup>3</sup> This is done based on difference-in-differences (DD)-estimators that compare changes in levels and gradients across cohorts and across states.

On average, the reform neither increased years of education nor the likelihood of being eligible to apply for university nor university graduation. We find, however, that the reform had a significant negative effect on the gradient in terms of years of education (which is measured as the time usually required to obtain the highest degree attained): the gradient in years of education decreased by about two-thirds to three-fourths of a year. This effect is entirely driven by males for which the gap in years of education decreased by 1.2 to 1.5 years, accounted for by an increase for males with uneducated parents by 0.40 to 0.45 years and a decrease for males with educated parents by about 0.6 years. We find that the overall effect is driven by both changes in the composition of males eligible for university and the composition of males that obtain tertiary degrees.

We rule out several alternative explanations for our finding by systematically changing the underlying sample of individuals. Importantly, we show that our results cannot be explained by systematic differences in levels and trends in pre-primary enrollment and differences in the pace and timing of Germany's educational expansion. Overall, our finding of a negative effect of the reform on the intergenerational transmission of education is remarkably robust to a number of

---

<sup>1</sup>For instance, countries differ in the age at which students are tracked and in whether tracking occurs within schools (i.e. sorting students into different classrooms as in the United States and Canada) or across schools (i.e. sending students to different types of schools as in some European countries).

<sup>2</sup>In German: *Orientierungsstufe*.

<sup>3</sup>We define this as the gap in educational outcomes between individuals with educated parents and those with uneducated parents.

tests we conduct.

The German reform we study offers a particularly well-suited setting to study the effects of tracking: education policies in Germany are to a large extent the responsibility of the states while the federal government is in charge of most other policy areas that may affect educational outcomes. Hence, our analysis is unlikely to be subject to confounding trends in unobserved socio-economic and institutional factors.<sup>4</sup> Because of Germany’s system of equalization payments between states and sharing of tax revenues (see [Bundesministerium der Finanzen, 2015](#)), there is limited potential for different trends in resources allocated to public education between states. Also, private education institutions that would potentially mitigate the effect of policy reforms aimed at improving intergenerational mobility played a negligible role in Germany’s education system during the time period we study. Finally, Germany stands out among OECD countries for early tracking at which large effects on educational careers may be expected.

Only a limited number of studies investigate the long-term effects of changes in the tracking regime.<sup>5</sup> Identification strategies that rely on cross-sectional variation (e.g. [Bauer and Riphahn, 2006](#)) potentially suffer from bias due to selection-on-unobservables or omitted variables ([Betts and Shkolnik, 2000](#); [Pischke and Manning, 2006](#); [Waldinger, 2007](#)). [Hanushek and Woessmann \(2006\)](#) exploit differences in the degree of tracking across countries based on an identification strategy that replaces changes within countries with changes across grades. They find a positive effect of tracking on performance inequality and no effect on performance levels. [Waldinger \(2007\)](#), however, presents evidence hinting at the presence of an omitted variable that affects both the intergenerational transmission of education and the likelihood of a country to implement early tracking.

[Brunello and Checchi \(2007\)](#) and [Schütz et al. \(2008\)](#) rely on cross-country variation in design features of education systems in order to estimate their impact on the importance of family background characteristics. The former authors focus on long-term outcomes such as earnings and employability and are thus closer to the present study. Both studies find that early tracking accentuates the importance of family background characteristics.

There are also several recent studies that investigate policy reforms within countries. [Malamud and Pop-Eleches \(2011\)](#) exploit a policy reform implemented in Romania in 1973 to study the effect of postponing tracking on educational outcomes. The reform increased significantly the proportion of students in general as opposed to vocational secondary schools. While they find a sharp increase in the proportion of students eligible to apply for university overall and among disadvantaged students, they do not find an increase in the likelihood of disadvantaged students to attend university. Their results are thus in stark contrast to ours.<sup>6</sup>

Studies similar to our own rely on variation across both cohorts and space to analyze de-tracking reforms in Scandinavia. [Meghir and Palme \(2005\)](#), [Pekkarinen et al. \(2009\)](#), and [Kerr](#)

---

<sup>4</sup>[Waldinger \(2007\)](#) argues that this may be an issue in cross-country comparisons such as the study by [Hanushek and Woessmann \(2006\)](#). See, however, [Woessmann \(2010\)](#).

<sup>5</sup>See [Betts \(2010\)](#) for a review.

<sup>6</sup>Note, however, that the setting also differs in many important ways.

et al. (2013) all report findings that are broadly in line with ours. Hall (2012), however, finds no effect of such a reform on university enrollment in Sweden. An important difference is that country-wide reforms in Scandinavia involved other significant changes to educational institutions and it is not possible to disentangle the effects of de-tracking from other components of these reforms.<sup>7</sup> Lower Saxony's reform, in contrast, entailed only adjustments to curricula in addition to delayed tracking. Nevertheless, our results below suggest that the pro-equality character of Germany's reform is broadly comparable to that of Scandinavian reforms.

The remainder of this paper is organized as follows. The next section describes the German education system and the process that finally led to the introduction of OS schools in Lower Saxony. Section 2 discusses channels through which tracking may affect educational outcomes and educational inequality. In section 4 we describe the data and our identification strategy. Our main results are presented in section 5 and robustness checks in section 6. Section 7 offers an interpretation of our results. In particular, we point towards possible explanations for the gender differences in effects that we find. Section 8 concludes.

## 2 Background: Germany's education system

### 2.1 Tracking in Germany's education system

Most states in Germany track students into three different types of secondary schools at the end of fourth grade when students are about ten years old:<sup>8</sup> low-achieving primary students usually attend the lower vocational track. The leaving certificate, awarded after five additional years of schooling, qualifies graduates to enroll in upper secondary vocational training courses (i.e. apprenticeships in the dual-system). Students with about average marks from primary education usually attend the intermediate secondary track. After another six years of schooling students are eligible to choose from an extended set of apprenticeships within the dual-system of vocational training. Enrollment in the academic track is recommended to highly-achieving students. This school type is the only secondary school track that awards after eight to nine years of schooling the *Abitur*, the most prestigious school-leaving certificate that permits students to apply to a university.

The streaming procedure varies across states but usually involves teachers formally recommending a secondary school track to parents based on their child's performance. In ten out of 16 states, parents have the final say about the placement of their child, whereas in the remaining six states, recommendations are binding yet parents have the right to let their child take an entry exam or attend test lessons.

---

<sup>7</sup>Sweden's reform entailed an increase in compulsory years of schooling and cash transfers to compensate families for foregone earnings. De-tracking in Finland coincided with the abolition of a vast network of private schools which were placed under municipal authority.

<sup>8</sup>Exceptions include Berlin and Brandenburg, in which elementary school comprises six grades, and Mecklenburg-West Pomerania, which tracks students after sixth grade but was part of the German Democratic Republic.

Tracks differ in several respects in terms of the quality and quantity of inputs (Dustmann et al., 2014): first, teachers in the academic track usually receive higher wages. Second, their university degree differs in terms of requirements, is more subject-oriented, and also typically takes one additional year to complete. Third, students in the academic track cover more topics and more advanced topics each year and they are often required to attend more hours per week. Finally, Brunello and Checchi (2007) report that the ratio of students to teachers in German schools varied substantially across tracks in 2004 with 11.89 students per teacher in the general track and 21.25 students per teacher in the vocational track.

Tracking across schools may be inconsequential if the education system exhibits a high level of permeability. In principle, German students are allowed to switch between tracks at any time if their academic records justify such a step. However, research on the topic suggests that switching between tracks prior to completion is rare. Mühlenweg (2008), for instance, examines administrative data from Hesse and reports for the school years 2003/2004 and 2005/2006 that more than 96 percent of students in grades five through seven remain in their initial track. Similar results are reported by Dustmann et al. (2014) who find that only two percent of students in the states of Bavaria and Hesse change track. Avenarius et al. (2001) present similar numbers for Lower Saxony. Moreover, Schnepf (2002) points out that students in the academic or upper vocational track are more likely to shift to the lower track than vice versa. On the other hand, upgrading upon completion of one of the two vocational tracks is quite common (e.g. Dustmann et al., 2014) and sometimes cited as evidence for the permeability of Germany’s education system. Students in the lower vocational track may switch to the upper vocational track or stay on for an additional year in order to obtain the school leaving-certificate awarded upon completion of the upper vocational track. Students in the upper vocational track, in turn, have the opportunity to continue schooling in the academic track if their academic record meets requirements.<sup>9</sup>

## 2.2 The introduction of OS schools in Lower Saxony

Serious attempts to reform Germany’s three-tiered education system were first evident during post-war decades when some states started experimenting with less stringent forms of tracking in a few selected schools. The goal at the time was to improve the selection of students into academic careers. However, efforts in this direction were terminated in Lower Saxony in 1964 (Schuchart, 2006). At the end of the 1950s, a federal advisory board formally recommended a prolongation of comprehensive schooling until sixth grade (Deutscher Ausschuss, 1959), a recommendation that had no effect on policies at the time (von Friedeburg, 1992). Only Hesse introduced schools that closely resembled these recommendations but would also retain schools in the traditional three-tiered system.

A follow-up body, the *German Education Council*, presented a blueprint for structural reforms

---

<sup>9</sup>Alternatively, they qualify to attend specialized academic track-schools that often have a special focus on a particular subject. Completion of such a specialized academic-track school allows them to apply to universities, although the range of subjects from which they may choose may be limited.

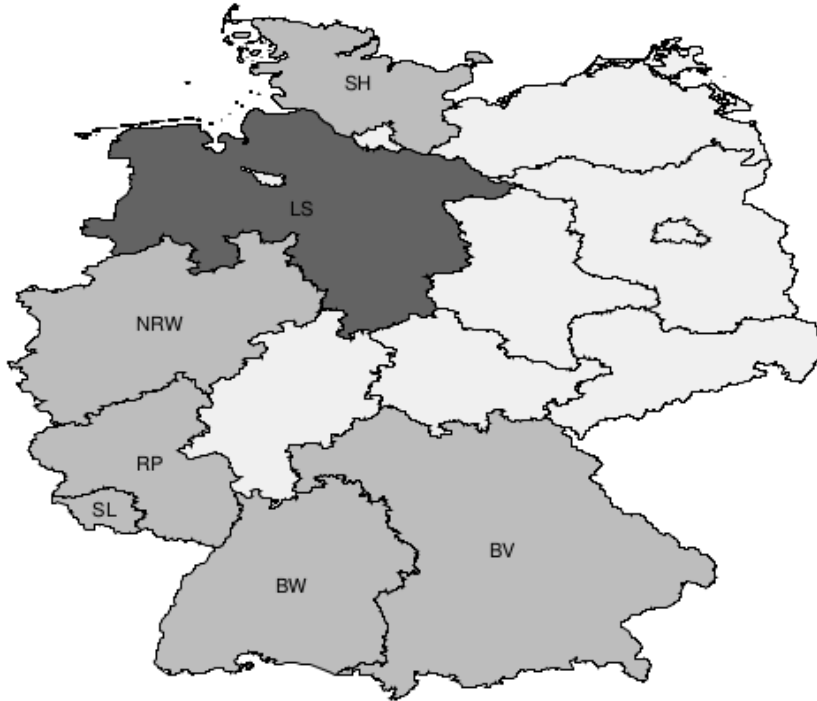


Figure 1: Treatment (dark gray) and control states (light gray). From North to South: SH (Schleswig-Holstein), LS (Lower Saxony), NRW (Northrhine-Westphalia), RP (Rhineland-Palatinate), SL (Saarland), BW (Baden-Württemberg), and BV (Bavaria).

that first mentioned the introduction of an *orientation stage* in 1970 (Deutscher Bildungsrat, 1970). In comparison to earlier plans, the focus was on tracking within schools across subjects and the dissemination of information about possible future careers to students and their parents (Schuchart, 2006). OS schools were to be completely independent of schools in the three-tiered system. In 1974, however, it became clear that this proposal would not be approved by a majority of states. While all states agreed in principle to changes to the school system, the compromise would leave the decision over whether or not to delay tracking to the states (Ziegenspeck, 2000, p. 81). While initial trials with OS-type schools were evident in several states, ultimately, only Lower Saxony and the city state of Bremen opted for tracking of all students at the age of twelve.<sup>10</sup>

Our analysis below compares changes across cohort in educational outcomes in students that received schooling in Lower Saxony to changes across cohorts that received schooling in other states of West Germany. We exclude the city states of Bremen, (West-)Berlin, and Hamburg as they differ in many important ways. We also exclude the state of Hesse which introduced an OS-style school but retained schools in the old system. The new school type was thus an

<sup>10</sup>Bremen established six years of primary education. However, from 1957 onward, there was an option to switch to the academic track already upon completing fourth grade (Schuchart, 2006, p. 70).

alternative, not a replacement. Figure 1 depicts the locations of the states we focus on within Germany.<sup>11</sup>

There are several aspects of the design of the OS in Lower Saxony over and above delayed tracking that are important for the present study: first, the administration of OS schools was independent of other secondary schools. This was not the case, for instance, in states such as Schleswig-Holstein and Rhineland-Palatinate in which OS schools were affiliated with secondary schools in the old three-tiered system in what amounts to an inconsequential re-labeling of fifth and sixth grade in the traditional system. Second, the reform entailed some degree of within-class tracking and ability grouping: teachers were expected to adjust tasks and demands within each class according to students' interests and abilities (Avenarius et al., 2001). Students were also grouped into three different levels of academic achievement during English and math lectures, where the three levels were to roughly reflect the demand levels of secondary school tracks (Ziegenspeck, 2000, pp. 262ff.). One could thus argue that the comparisons that we rely on in our analysis below are between tracking across schools and a system with a strong element of tracking within schools. Third, OS schools employed teachers from all secondary school tracks on a part-time basis. As every class was to be taught by teachers of all secondary school tracks, students were exposed to teachers with profound knowledge of the everyday practice at the secondary schools in order to ensure a suitable tracking. Fourth, teachers at OS schools were asked to regularly document students' behavior and academic progress according to fixed criteria. This was aimed at nudging teachers towards a more objective judgment of children's academic potential (Hornich, 1976, p.110).

Finally, there was an emphasis on frequent consultations with parents in order to inform them about their child's future perspectives and to explain and to convince them of the ensuing recommendation. This was important since from 1979 onward, parents had the final say about whether or not to heed the schools' recommendation.<sup>12</sup> To summarize, the reform intended to balance the advantages associated with creating homogeneous groups with an effort to dilute the influence of parental background.

---

<sup>11</sup>Out of the six remaining states that will serve as controls in our analysis below, two, Northrhine-Westphalia and Saarland, did not introduce OS schools of any stripe by 1975. In both states, however, the introduction was planned for the second half of the 1970s (Haenisch and Ziegenspeck, 1977, pp. 40ff). Ultimately, reform efforts would run out of steam or would be prevented through referenda (Rösner, 1981). While there were experiments with OS-type schools in the two southern states of Bavaria and Baden-Württemberg by 1975, these states eventually still opted for tracking after fourth grade. For instance, Bavaria would delay tracking into the two lower tracks by two years until the early 2000s but tracking after fourth grade into academic and vocational tracks was continued Piopiunik (2013, e.g.). Rhineland-Palatinate and Schleswig-Holstein re-labeled fifth and sixth grade (as did Northrhine-Westphalia eventually) but this did not lead to changes in the tracking regime. In Lower Saxony, OS schools were the norm until they were abolished in 2004.

<sup>12</sup>There is anecdotal evidence that parents at the time had a tendency to overrule recommendations in favor of choices typical of their social class. There was also an understanding among educators that working class parents had to be convinced that their child did not necessarily have to attend one of the two vocational tracks (Ziegenspeck, 2000, pp. 146ff.).



## 3 Conceptual framework

### 3.1 Peer effects

Tracking may affect the efficiency of educational production as well as the variation in outcomes through altering the relationship between individual and peer quality.<sup>13</sup> While the early empirical literature on direct peer effects is beset with econometric problems (e.g. Manski, 1993; Sacerdote, 2001), recent studies suggest that students benefit from the presence of higher-achieving peers (see, *inter alia*, Sacerdote, 2001; Ding and Lehrer, 2007; Lavy et al., 2011). Hence, if tracking succeeds in matching peers of similar quality, one would expect an increase in the overall variation in achievement. Moreover, if own achievement is linked to parental education, one would expect to find a positive relationship between tracking and the gradient between parental education and own education.

At the same time, it is often argued that tracking may allow for efficiency gains in educational production. Tracking will result in more homogeneous groups which, in turn, may allow teachers to tailor lessons more specifically to students' needs. Tracking may thus benefit *all* students. However, empirical studies on this channel are ambiguous (e.g. Epple and Romano, 2011). While Ding and Lehrer (2007) and Duflo et al. (2011) find a positive effect of increasing peer homogeneity on achievement, Lyle (2009), for instance, finds that a *higher* variance in peer math scores benefits students.

### 3.2 Mis-allocation of students to tracks

A second point relates to the precision with which students' are tracked. While track choice should arguably be based on academic potential, it is often maintained that this is difficult to observe initially and that the signal becomes stronger over time (e.g. Brunello et al., 2007). Early tracking may thus be associated with a mis-allocation of students to tracks. In particular, it may be the case that non-cognitive skills such as attentiveness become more important for track choice at an early age when cognitive skills are still difficult to observe. Non-cognitive skills, in turn, have been shown to be related to parental background variables such as parents' educational attainment (Segal, 2008; DiPrete and Jennings, 2012).

There is indeed evidence that early tracking fails in separating students effectively by academic potential, particularly in Germany. First, there is considerable overlap in test scores between different school tracks (Baumert et al., 2003). Lehmann and Peek (1997) find that students of uneducated parents have to score significantly higher in standardized tests in order to receive an academic track recommendation with the same probability as students with educated parents.

Second, track choice is often found to be determined by variables that are arguably unrelated to academic potential. A large literature, for instance, documents relative age-effects in education: Puhani and Weber (2007), Mühlenweg and Puhani (2010), Jürges and Schneider (2011),

---

<sup>13</sup>Epple and Romano (2011) provide a comprehensive review of the literature on peer effects in education.

and [Dustmann et al. \(2014\)](#) all find that students’ exact birthday relative to an arbitrary enrollment cut-off date predicts track choice in Germany.<sup>14</sup> While such findings are indicative of inefficiencies associated with tracking, both [Jürges and Schneider \(2011\)](#) and [Dustmann et al. \(2014\)](#) find no evidence for a persistent effect of relative age on educational outcomes. Moreover, [Jürges and Schneider \(2011\)](#) find no evidence that the age at which states track affects the strength of the relative age effect based on variation across states.

On the other hand, initial misallocation may be inconsequential in a system such as Germany’s in which upgrading upon completion is common (see section 2.1). In a recent study, [Dustmann et al. \(2014\)](#) use relative age at birth as an instrument to show that track choice is largely inconsequential for marginal students. They argue that built-in flexibilities, the possibility of upgrading upon completion, eventually allow marginal students in the lower tracks to fully compensate for the exposure to low-ability peers.

There is also an increasing interest in gender differences in student achievement, the evolution of such differences over time, and their interaction with tracking. [Bedard and Cho \(2010\)](#) report that tracking is pro-female in Germany in that females are placed in classes with higher average ability. Both [Lehmann and Peek \(1997\)](#) and [Jürges and Schneider \(2011\)](#) report that boys are less likely to be recommended to the academic track in Germany *conditional on academic achievement*, suggesting that girls outperform boys in other relevant dimensions.<sup>15</sup> Interestingly, [Jürges and Schneider](#) find no evidence that the gender effect varies across states with different tracking procedures and while the relative age-effect seems to fade over time, the number of female students in the academic track is still greater in ninth grade. The authors conclude that delaying tracking by two years would not reduce gender bias in track attendance.

## 4 Estimating the impact of the reform on educational outcomes

### 4.1 Data and descriptives

In this section, we describe some key variables of our analysis and explain coding decisions. The dataset we use is the German Socio-Economic Panel (GSOEP v29), a nationally representative survey carried out on an annual basis between 1984 and 2012. See [Wagner et al. \(2007\)](#) for a description of the dataset.

Outcome variables in our analysis are total years of education and two binary variables indicating whether an individual has attained the highest school-leaving certificate (i.e. is eligible to apply to a university without restrictions) and whether an individual has graduated from a university. Means are reported in panel A of table 1. ‘Years of education’ in this dataset refer to the number of years usually required to obtain certain degrees, not to the time spent in education,

<sup>14</sup>[Schneeweis and Zweimüller \(2014\)](#) present evidence for Austria in which tracking also occurs at the age of ten.

<sup>15</sup>[Jürges and Schneider](#) relate this finding to “differences in verbal and non-cognitive skills at age 10.”

Table 1: Means of key variables by treatment status and gender, entire estimation sample.

	Other states			Lower Saxony		
	All	Females	Males	All	Females	Males
<i>Panel A. Outcomes</i>						
Years of education	12.06	11.83	12.30	12.19	11.98	12.41
Eligible for university	0.32	0.27	0.38	0.33	0.27	0.38
University degree	0.17	0.13	0.22	0.18	0.13	0.24
<i>Panel B. Individual background characteristics</i>						
<i>Panel B1. Basic characteristics</i>						
Year of birth	1958	1958	1958	1958	1958	1958
Educated parent	0.37	0.36	0.37	0.37	0.37	0.38
<i>Panel B2. Childhood place of residence was mostly...</i>						
...city.	0.21	0.20	0.21	0.12	0.11	0.12
...large town.	0.17	0.17	0.17	0.16	0.15	0.17
...small town.	0.21	0.22	0.21	0.22	0.22	0.22
...rural area.	0.36	0.36	0.36	0.46	0.48	0.44
<i>Panel B3. Migratory background</i>						
Direct	0.04	0.04	0.05	0.02	0.01	0.02
Indirect	0.05	0.05	0.05	0.03	0.02	0.03
Unspecified	0.00	0.00	0.00	0.00	0.00	0.00
No migratory background	0.91	0.91	0.90	0.96	0.96	0.95
<i>Panel B4. Sample information</i>						
Last observed in 2012	0.62	0.62	0.62	0.62	0.64	0.61
Observations	1,593	2,430	1,977	3,484	981	488

Statecodes as in notes to figure 1. Based on GSOEP v29-data.

and includes all stages from primary to tertiary education.<sup>16</sup> Grade repetition will thus not be reflected in this measure of educational attainment.

Variables that serve as controls in our analysis capture socio-demographic characteristics such as year of birth<sup>17</sup> and gender. We also always include a complete set of indicators for the size of the respondent's locality during childhood<sup>18</sup> and migrant status.<sup>19</sup>

Information on parents' educational attainment is available for almost all individuals in the data. There are two variables for each mother and father that relate to the school-leaving certificate and the type of tertiary schooling or vocational training completed. We code a binary variable equal to unity if either the mother or father has (i) attained the highest school-leaving certificate, has (ii) completed the upper vocational track of secondary *and* a full course vocational

<sup>16</sup>The maximum this variable takes is 18 years, 13 years until the highest school-leaving certificate plus five years in order to obtain a university degree. Obtaining a Bachelor's degree or a degree from a university of applied sciences usually takes three years in Germany. Depending on the type of job, vocational training adds 1.5 or two years to the total. Completing the lower and upper secondary tracks takes nine and ten years, respectively.

<sup>17</sup>Since the school year starts after the summer, we re-code this variable to reflect schooling cohort.

<sup>18</sup>Categories are 'no information available', 'city', 'large town', 'small town', and 'rural area'.

<sup>19</sup>Categories are 'no information available,' 'no migratory background', 'direct migratory background', 'indirect migratory background', and 'migratory background (not further specified)'.

education, or has (iii) completed training as a clerk, a public health worker, a civil servant, or an engineer, or holds a degree from a tertiary education institution.<sup>20</sup> This results in about 35 percent of our observations being classified as having educated parents (see panel B1 of table 1). Note that our definition allows for either the father or the mother (or both) to have attained this level of education. Hence, the presence of only one educated parent is assumed sufficient to generate the relevant externality at the household-level (Basu and Foster, 1998).

Logit regressions reveal that our indicator of parents' educational attainment is highly predictive of parents' type of occupation<sup>21</sup> and the likelihood that respondents engaged in extracurricular activities at the age of 15 (such as having actively played a musical instrument or having done sports).<sup>22</sup> To further investigate differences between educated and uneducated parents, we also correlate this variable with an indicator of parents' preoccupation with respondents' academic achievement. This variable is coded on a five-point scale ranging from 'very much' to 'not at all'. We code the highest two levels as a binary variable. Results from logit regression in which we condition on own year of birth, parents' years of birth, parents' age, and the state of the last school visit indicate that having educated parents increases the odds that they were at least 'rather preoccupied' by a factor greater than two for both males and females.

The GSOEP data do not provide direct information about the type of school individuals attended at the age of ten. We therefore supplement the data with information from the statistical reports on schooling in Lower Saxony (Landesamt für Statistik Niedersachsen). These reports tabulate the number of students by grade and birth year in different school types. Combining these data allows us to calculate the percentage of students of one cohort that attended an OS school in a given school year.<sup>23</sup> Enrollment at the time occurred at the age of six and school years start after the end of the summer in July or August. For instance, an individual born in the first half of 1960 was supposed to start schooling in school year 1966/1967 and somebody born in the second half of that year was supposed to start schooling in school year 1967/1968. We therefore associate individuals of a given cohort in Lower Saxony with the share of students in OS schools with one of two subsequent school years depending on whether the individual was born during the first or second half of the year. Figure 2 plots the variable  $OS_c$  against birth years for individuals that attended schools in Lower Saxony.

There is usually no direct information in the GSOEP dataset about the state in which an individual resided at the age of ten. A second challenge is thus to infer the state in which individuals went to school at that age. Clearly, imputing the current state would confound treatment and control due to inner-German migration. There are several variables in the dataset that allow us to close in on the required information such as whether the current place of residence

<sup>20</sup>This could be either a regular university (including foreign institutions) or a university of applied sciences.

<sup>21</sup>Conditional on fathers' (mothers') year of birth and own year of birth, we find having an educated father (mother) increases the odds of having a father (mother) that worked as a skilled, white-collar worker by the time respondents were 15 years old by a factor greater than four (three). These estimates are highly significant.

<sup>22</sup>Conditional on own year of birth and parents' years of birth, having educated parents increases the odds of having played an instrument and having done sports by a factor of about 2.5 and 2.3 for boys, and 2.5 and 1.4 for girls, respectively.

<sup>23</sup>We ignore private schools which play a negligible role in Germany.

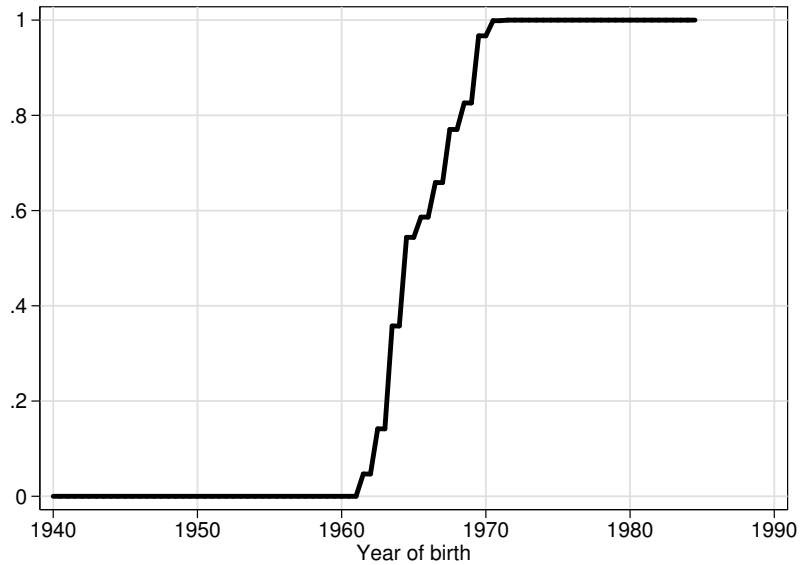


Figure 2: OS share against birth year in Lower Saxony.

was also the childhood place of residence and the state of the last school visit. We also drop observations on individuals that did not attend school in one of the West German states. The exact procedure is described in appendix A. While these steps are reasonable, there will still be misclassification error. Note, however, that confounding treatment and control will tend to bias coefficients towards zero. Our estimates may thus be interpreted as lower bounds of the true effect.

Since our interest is exclusively in variables that do not change over time, we discard all observations except the most recent for each individual. This is 2012 for more than half of the observations in our sample but sometimes an earlier year (see panel B4 of table 1).

As was discussed in section 2, we retain only observations on individuals that are likely to have received their schooling in one of seven states. Due to low case counts and associated privacy issues, two states, Rhineland-Palatinate and Saarland, were treated as one entity in the dataset. We further retain only observations on individuals above the age of 28 as these have a high likelihood of having completed their education.<sup>24</sup> In consequence, all individuals in our sample were born before 1985. We also exclude individuals born before 1950.

<sup>24</sup>Schooling usually starts at the age of six and completing secondary education with the *Abitur*, the highest school-leaving certificate in Germany, requires 13 years of schooling over the time period that we study. Mandatory military service never exceeded two years. If we add five years required to obtain a university degree, the age at which one would complete university is 26.

## 4.2 Empirical specifications

We estimate the causal effect of the reform on educational outcomes and the intergenerational transmission of education. We conduct this analysis separately for males and females. Our analysis is retrospective: individuals are observed only after they have acquired their education. We employ a DD estimator with a first difference across cohorts and a second difference across individuals in different states. Since the school year starts after the summer, a cohort  $c$  comprises all individuals born during the second half of a given year or during the first half of the subsequent year.

Define the treatment variable as the product of the percentage of a cohort in OS schools in Lower Saxony and an indicator variable for Lower Saxony, i.e.  $D_{sc} = OS_c \times LS_s$ . Note that  $D_{sc}$  is not a binary variable but takes on a range of values between zero and unity for cohorts born in the 1960s. A standard estimator for the effect of the reform on educational outcomes is

$$y_{isc} = \alpha D_{sc} + \mathbf{x}'_{isc} \gamma + \lambda_s^b + \tau_c + \epsilon_{isc}, \quad (4.1)$$

where  $y_{isc}$  is the educational outcome for individual  $i$  who went to school in state  $s$  and is a member of cohort  $c$ .  $y_{isc}$  denotes years of education (as defined above) or binary indicators of university eligibility and graduation from university. For the latter two, the model is a linear probability model (LPM).<sup>25</sup>  $\alpha$  is interpreted as the effect of the reform on the respective outcome.

$\mathbf{x}_{isc}$  is a matrix of time-invariant, individual characteristics. These include indicators for the respondent's gender, migrant status, and type of locality in which individuals spent their childhood. Additional checks that we report on in the next section suggest that there may have been some variation in the gradient between parental education and own education across states prior to the reform. We therefore always include state-background-fixed effects  $\lambda_s^b$ , where  $b \in \{low, high\}$  indicates parental education.  $\tau_c$  denotes cohort-fixed effects. When the sample is pooled across males and females, we interact  $\lambda_s^b$ ,  $\tau_c$ , and all variables in  $\mathbf{x}_{isc}$  with an indicator of respondents' gender. Finally,  $\epsilon_{isc}$  is the usual white noise-error term.

A variation of (4.1) that also allows for background-specific cohort effects is

$$y_{isc} = \alpha D_{sc} + \mathbf{x}'_{isc} \gamma + \lambda_s^b + \psi_c^b + \epsilon_{isc}, \quad (4.2)$$

where  $\psi_c^b$  denotes a set of cohort-background-fixed effects.

We hypothesize that the effect of the reform on own education depends on parental education, that is, that the reform had an effect on the gradient between parental education and own education. A straight-forward way to test this is to include an interaction between the treatment variable and a binary indicator of having educated parents. The model can then be written as

$$y_{isc} = \beta_1 (D_{sc} \times I[b = high]) + \beta_2 D_{sc} + \mathbf{x}'_{isc} \gamma + \lambda_s^b + \tau_c + \epsilon_{isc}. \quad (4.3)$$

---

<sup>25</sup>Estimating probit and logit models does not alter our results. Note also that the models discussed here are saturated in their main part, that is, we compare differences in means that are bounded between zero and one.

The parameters of interest here are  $\beta_1$  and  $\beta_2$ . The former measures the effect of the reform on the gradient and the latter the effect on individuals with uneducated parents. The effect on individuals with educated parents is calculated as the sum over both coefficients. In order to obtain an estimate of the average effect of the reform, we also present results from estimating (4.3) and similar models without the first term on the right hand-side.

The reform was implemented during a time of a nation-wide educational expansion, mostly a response to rapid population growth during the 1960s and 1970s. We therefore estimate an alternative model that allows for differences in trends in educational attainment across states by including state-cohort-fixed effects (denoted  $\phi_{sc}$ ):

$$y_{isc} = \beta_1(D_{sc} \times I[b = high]) + \mathbf{x}'_{isc}\gamma + \lambda_s^b + \phi_{sc} + \epsilon_{isc}. \quad (4.4)$$

Note that the main effect of the reform is no longer identified and thus excluded here.  $\beta_1$ , however, is still identified from within-cohort, within-state variation.

The identifying assumption for the causal effect of the reform on educational outcomes in the framework above is that trends in outcomes would not have systematically differed between Lower Saxony and other states in the absence of the reform. While this may be reasonable, the assumption required to obtain an unbiased estimate of the effect of the reform on the gradient from (4.4) is more demanding as it required that the gradient remains constant over time except for the effect of the reform. We relax the latter assumption by including  $\psi_c^b$  to allow for common changes in the gradient across cohorts:

$$y_{isc} = \beta_1(D_{sc} \times I[b = high]) + \mathbf{x}'_{isc}\gamma + \lambda_s^b + \psi_c^b + \phi_{sc} + \epsilon_{isc}. \quad (4.5)$$

In this specification, the assumption required to obtain a unbiased estimate of the effect of the reform on the gradient is that there would not have been any systematic differences in trends *in the gradient* absent the reform.

A general concern one may have in estimating the above specifications is that the reform may have caused selective migration: if families decided to move from or to Lower Saxony in response to the reform and if this decision was correlated with parental education, our estimates would be biased. There are, however, no reports of such responses that we know of. Note also that this would have required parents to move to another state which, in most cases, would have required them to leave their job. It seems implausible that a reform that only affected two grades would have induced this behavior.

If there had been a significant migratory response, however, the reform would be associated with a change in the state-specific probability of having educated parents. To check that this is not the case, we regressed the binary indicator of having educated parents on the treatment variable as well as a complete set of state- and cohort-fixed effects (results not reported). We also included all controls that we include when we estimate (4.3) and ran this regression for the pooled sample and males and females separately. We found no evidence that the reform

had an effect on the probability of having educated parents: all coefficients were close to zero and statistically insignificant. We conclude that selection due to strategic migration is not a confounding element in our analysis.

### 4.3 The reform as a natural experiment

Before we turn to our results, we investigate (1) whether trends in educational outcomes and gradients differed prior to the implementation of the reform and (2) whether raw gradients differed between states prior.

#### 4.3.1 Pre-existing differences in trends in outcomes and gradients

Our identification assumption for the effect of the reform on educational outcomes and the gradient is the usual common trends-assumption, that is, we assume that trends in outcomes and gradients would not have differed systematically across states in the absence of the reform. While this cannot be tested directly, a finding of common trends *prior to the reform* would suggest that this assumption is plausible.

To test for the existence of systematic differences in pre-reform trends in outcomes and gradients, we restrict the sample to individuals born before 1961 and in the first half of that year. These cohorts were not yet affected by the reform. We then run two specifications based on equations (4.2) and (4.5) above:

$$y_{isc} = \xi_1(LS_s \times c) + \mathbf{x}'_{isc}\gamma + \lambda_s^b + \tau_c + \epsilon_{isc} \quad (4.6)$$

and

$$y_{isc} = \xi_2(LS_s \times I[b = high] \times c) + \mathbf{x}'_{isc}\gamma + \lambda_s^b + \psi_c^b + \phi_{sc} + \epsilon_{isc}, \quad (4.7)$$

where all variables are defined as before. Testing whether our estimates of  $\xi$  are zero in equation (4.6) and (4.7) will amount to testing the null of no pre-existing trends in outcomes and gradients, respectively. A positive, significant estimate for  $\xi_1$  in (4.6) would suggest that educational attainment was increasing at a higher rate in the treatment state. Similarly, a positive and significant estimate for  $\xi_2$  in (4.7) would indicate that the gradient between own education and parental education was increasing at a higher rate in the treatment state. Note that the coefficient on  $(LS \times c)$  is not identified in (4.7) as we include state-cohort-fixed effects.

While using a linear model to estimate the treatment effect in a DD specification when the outcome is a binary variable is innocuous, including linear cohort-trends will result in these models being miss-specified. We therefore also report results from estimating probit-versions of (4.6) and (4.7).

Results from estimating equations (4.6) and (4.7) are reported in panels A and B of table 2. Standard errors clustered at the state-cohort-level are reported in parentheses in all tables



Table 2: Tests for pre-existing trends in outcomes and gradients by gender.

	Females					Males				
	Years of education		University eligibility		University graduation	Years of education		University eligibility		University graduation
	OLS (1)	OLS (2)	OLS (3)	Probit (4)	Probit (5)	OLS (6)	OLS (7)	OLS (8)	OLS (9)	Probit (10)
<i>Panel A: Testing for common trends in outcomes prior to the reform.</i>										
$LS_s \times c$	0.005 (0.013)	0.003 (0.002)	0.015 (0.011)	0.001 (0.002)	0.008 (0.009)	0.002 (0.016)	0.001 (0.003)	0.004 (0.011)	0.000 (0.003)	0.000 (0.013)
Observations	3,268	3,268	3,265	3,268	3,265	3,107	3,107	3,104	3,107	3,107
R-squared	0.171	0.131		0.106		0.153	0.100		0.103	
<i>Panel B: Testing for common trends in gradients prior to the reform.</i>										
$LS_s \times I[b = high] \times c$	0.003 (0.044)	0.005 (0.007)	0.005 (0.024)	0.001 (0.006)	0.002 (0.022)	0.041 (0.043)	0.009 (0.006)	0.025 (0.017)	0.014* (0.008)	0.049* (0.027)
Observations	3,268	3,268	3,206	3,268	3,013	3,107	3,107	3,092	3,107	3,078
R-squared	0.206	0.164		0.147		0.189	0.135		0.146	

Standard errors clustered at the state-cohort-level in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. All regressions include dummies for the size of respondents' childhood place of residence and migrant status. When we use the pooled sample, we also include a dummy variable for the respondent's sex and interact it with all other controls (including fixed effects). Based on GSOEP v29-data.

Table 3: Variation in raw gradients in years of education prior to the reform (1940–1960 birth cohorts).

	Females			Males		
	(1)	Mothers (2a)	Fathers (2b)	(3)	Mothers (4a)	Fathers (4b)
Educated parents	2.11*** (0.21)	1.14*** (0.38)	1.74*** (0.33)	2.58*** (0.32)	2.28*** (0.37)	1.71*** (0.46)
Educated parents × BV	-0.44 (0.33)	0.50 (0.50)	-0.86** (0.43)	-0.84** (0.39)	-1.28** (0.56)	-0.54 (0.60)
Educated parents × BW	0.08 (0.37)	0.73 (0.55)	-0.46 (0.45)	-0.36 (0.37)	-1.59*** (0.51)	0.02 (0.61)
Educated parents × NRW	-0.01 (0.30)	0.14 (0.47)	-0.12 (0.41)	-0.60 (0.37)	-0.52 (0.50)	-0.18 (0.55)
Educated parents × RP	-0.46 (0.45)	-0.25 (0.56)	-0.11 (0.55)	-0.97* (0.53)	-1.39** (0.59)	-0.05 (0.78)
Educated parents × SH	-0.50 (0.43)	-0.55 (0.68)	-0.14 (0.66)	-0.96* (0.49)	-0.55 (0.94)	-0.17 (1.05)
<i>Joint significance of interactions.</i>						
<i>p</i> -value	0.51	0.32	0.32	0.18	0.03	0.93
Observations	3,268	3,268	3,268	3,107	1,680	1,680
R-squared	0.13	0.15	0.15	0.11	0.15	0.15

Standard errors clustered at the state-cohort-level in parentheses. Coefficient estimates reported in columns (2a) and (2b) and those in (4a) and (4b) pertain to the same models. Lower Saxony is the excluded category. Statecodes as in notes to figure 1. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. Based on GSOEP v29-data.

that follow.<sup>26</sup> There is no evidence that pre-reform trends in outcomes in Lower Saxony differed systematically from those in other states: estimated coefficients are close to zero and statistically insignificant. There is also no evidence that trends in the gradient differed across states for females. For males, we find a positive albeit insignificant coefficient on the three-way-interaction term when we test for pre-reform trends in the gradient in years of education, suggesting that the gap may have widened: however, the point estimate suggests that the gradient increased over ten years by less than half a year in Lower Saxony *vis-à-vis* other states. The coefficients in columns (9) and (10) are also positive and even significant at the five-percent-level. Note, however, that this suggests that had these trends continued in the absence of the reform, we would obtain a *downward-biased* estimate of the effect of the reform on the gradient. Since our results below show that the gap narrowed, it may be interpreted as a lower bound for the effect of de-tracking on the gradient.

### 4.3.2 Pre-existing differences in gradients

Next, we investigate pre-reform differences in raw gradients. Table 3 reports results from regressing years of education<sup>27</sup> on parental education and parental education interacted with state-dummies for the subset of males (panel A) and females (panel B) that turned ten before the onset of the reform. Lower Saxony is the reference category so that coefficients on interaction terms are estimated differences with the gradient observed in that state.

We first focus on differences in gradients for males. The coefficient on the main effect in column (1) is highly significant and suggests that males with educated parents in Lower Saxony have on average about 2.6 additional years of education. There is no evidence for systematic differences in gradients at this point: the  $p$ -value from an  $F$ -test for the null that all coefficients on interactions are jointly zero is 0.43. The estimate of the pre-reform gradient is significantly lower only in Rhineland-Palatinate yet the point estimate suggests a much-reduced importance of parental education. It is also interesting to see that the treatment state has had the largest gradient of all states.

To further investigate this, we estimate a model that includes fathers' and mothers' education separately as well as the respective state-interactions. While this is only one empirical model, results are reported separately for mothers' and fathers' education in columns (3a) and (3b), respectively. This shows that the overall differences are mainly driven by differences in the effect of maternal education on own education. Turning to panel B, we see that differences are similar qualitatively yet less pronounced for females and mostly insignificant. Fathers' education may have played a larger role in Baden-Württemberg.

Differences in the gradient prior to the reform are not *per se* a threat to our identification strategy. If these differences result from unobserved, time-invariant heterogeneity, they will be absorbed by state-background-fixed effects. It may be the case, however, that low levels of inter-generational mobility in Lower Saxony were perceived as a problem prior to the reform and that this triggered—in addition to the reform—further policy responses. In line with the notion that mothers are the primary care-givers in the absence of childcare, we show in a supplementary analysis in appendix B that persistent differences in pre-primary enrollment rates explain pre-reform differences in the gradient. These differences, in turn, are addressed in a robustness check in section 6 below in which we exclude states that differed in terms of pre-primary enrollment rates and, (at the same time, differences in pre-reform gradients.

Table 4: Impact of the reform on years of education by gender.

	Pooled		Females		Males	
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	-0.011 (0.131)	-0.032 (0.133)	0.013 (0.168)	-0.006 (0.169)	-0.039 (0.207)	-0.061 (0.206)
Observations	12,810	12,810	6,630	6,630	6,180	6,180
R-squared	0.176	0.184	0.189	0.196	0.155	0.164
<i>Fixed effects:</i>						
State-background	✓	✓	✓	✓	✓	✓
Cohort	✓		✓		✓	
Cohort-background		✓		✓		✓

Standard errors clustered at the state-cohort-level in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. All regressions include dummies for the size of respondents' childhood place of residence and migrant status. When we use the pooled sample, we also include a dummy variable for the respondent's sex and interact it with all other controls (including fixed effects). Based on GSOEP v29-data.

## 5 Results

We now turn to the main results of our analysis. Coefficient estimates reported in column (1), (3), and (5) of table 4 result from estimating (4.1) with the pooled sample and for females and males, respectively. Columns (2), (4), and (6) report analogous results from estimating (4.2), that is, the specification including background-specific trends over cohorts. Both specifications result in points estimates that are close to zero and insignificant at conventional levels. The reform does not seem to have had an effect on years of education *on average*. In other words, we find no evidence that tracking is associated with an increase in educational attainment.

Next, we investigate whether the reform has had an effect on the intergenerational transmission of education. Results from estimating specifications (4.3)–(4.5) are reported in columns (1)–(3) of table 3. Results for the pooled sample are reported in panel A and results for females and males are reported in panels B and C, respectively.

Results in column (1) of panel A suggest that the effect of the reform on years of education did depend on parental background. The reform is associated with an increase in years of education by one-fourth of a year for individuals with uneducated parents. However, the estimate is not significantly different from zero ( $p$ -value = 0.147). We do find, however, that the effect on individuals with educated parents was negative and significantly different from zero: the point

<sup>26</sup>We also experimented with conventional, robust ('sandwich'-type) standard errors, and standard errors clustered at the level of cohorts and states. Clustering at the state-cohort-level turned out to be the most conservative option and not very different from conventional standard errors, robust standard errors, and standard errors clustered at the level of cohort. Clustering at the state-level resulted in much smaller standard errors. We address potential problems with standard errors in a robustness check in section 6.1 below.

<sup>27</sup>Results for university eligibility and university graduation are similar.

Table 5: Impact of the reform on gradient in years of education, university eligibility, and university graduation by gender.

	Years of education			University eligibility			University graduation		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)		
<i>Panel A. Pooled sample (N = 14, 022).</i>									
Ref. × ed. parents	-0.486* (0.271)	-0.575* (0.301)	-0.886*** (0.320)	-0.043 (0.060)	-0.090 (0.059)	-0.108** (0.052)	-0.150*** (0.054)		
Reform	0.172 (0.172)								
R-squared	0.176	0.202	0.210	0.161	0.167	0.142	0.150		
<i>Panel B. Females (N = 7, 236).</i>									
Ref. × ed. parents	-0.074 (0.332)	-0.056 (0.351)	-0.332 (0.385)	0.042 (0.069)	0.010 (0.071)	-0.042 (0.060)	-0.090 (0.066)		
Reform	0.042 (0.209)								
R-squared	0.189	0.215	0.222	0.167	0.173	0.137	0.145		
<i>Panel C. Males (N = 6, 786).</i>									
Ref. × ed. parents	-0.973* (0.509)	-1.163** (0.510)	-1.509*** (0.541)	-0.140* (0.083)	-0.202** (0.086)	-0.183** (0.090)	-0.217** (0.094)		
Reform	0.314 (0.259)								
R-squared	0.156	0.183	0.192	0.131	0.137	0.130	0.137		
<i>Fixed effects:</i>									
State-background	✓	✓	✓	✓	✓	✓	✓		
Cohort	✓								
State-cohort		✓	✓	✓	✓	✓	✓		
Cohort-background			✓		✓	✓	✓		

Standard errors clustered at the state-cohort-level in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. All regressions include full sets of dummy variables for the size of respondents' childhood place of residence and migrant status. When we use the pooled sample (panel A), we also include a dummy variable for the respondent's sex and interact it with all other controls (including fixed effects). Based on GSOEP v29-data.

estimate here is  $\widehat{\beta}_1 + \widehat{\beta}_2 = -0.358$  and we reject that this parameter is zero at the ten-percent level ( $p$ -value = 0.058). The estimate on the interaction terms is  $-0.616$  and significant at the five-percent level, suggesting that the gradient between parental and own education in terms of effective years of education decreased by about three-fifths of a year.

We find similar estimates for the effect on the gradient when we include state-cohort-fixed effects (column (2)) and after allowing for common shifts in the gradient over time (column (3)). Estimates in columns (4)–(7) show that the effect on effective years of education is driven by both a change in the composition of university eligibles and graduates. For instance, the estimates in columns (6) and (7) suggest that the gap in the probability of obtaining a university degree decreased by more than ten percentage points. However, the former change is not significantly different from zero.

Comparing across panels we see that the overall effect is driven almost entirely by males. While we often find a negative sign on the coefficient that multiplies the interaction effect for females, the effect is small and statistically insignificant. For males, on the other hand, we find that the gap in years of education decreased by between one and 1.3 years and that all estimates in this case are significant at the five-percent level.

## 6 Robustness

In this section we conduct different robustness checks that address potential problems with our analysis. We first investigate the robustness of our results to aggregating data at the level of state-cohorts. We then investigate whether our results are robust to alternative sample restrictions that allow us to rule out other potential explanations for our findings. Finally, we present results from a placebo test.

### 6.1 Accounting for grouped data

The previous section shows that the result of a decrease in the gradient for males is robust to several alternative estimation methods. One concern may be that the treatment variable varies only at the state-cohort-level. Since the number of clusters is limited, standard errors may be unreliable. Following [Angrist and Pischke \(2009, p. 313\)](#), we collapse our dataset at the state-cohort-level. We then compute the mean difference in years of education between individuals with educated and those with uneducated parents. We regress this gap for males and females separately on the reform variable and include both cohort- and state-fixed effects. We also include variables that capture the share of individuals with direct and indirect migratory background and four variables indicating the level of urbanization of the childhood place of residence. We estimate these models by weighted least squares using the group size as weights.

There are two advantages of this approach: first, it makes our identifying assumption with respect to the reform’s effect on the gradient explicit: the gradient is the outcome variable in a

Table 6: DD estimates of the reform’s effect on gradient in years of education based on grouped data.

	Females (1)	Males (2)
Reform	-0.354 (0.448)	-1.554*** (0.508)
R-squared	0.217	0.261
Observations	266	268

Asymptotic standard errors in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. Observations weighted by the number of individuals in each cohort-state. All regressions include sets of cohort- and state-fixed effects. We also include the share of individuals with direct and indirect migratory background, and the share of individuals that grew up in rural areas, small towns, large towns, and cities in each cohort-state-cell as additional controls. Based on GSOEP v29-data.

standard difference-in-differences equation. Second, it is also evident this way that the asymptotics of our analysis are based on the number of cohort-state-groups. Because the group means are asymptotically normally distributed, the finite-sample properties should closely resemble those of a regression with normally distributed errors. Usual asymptotic standard errors are therefore consistent given the grouped structure in the data (Angrist and Pischke, 2009, p. 313).

Results reported in table 6 confirm our previous findings: for males, the introduction of OS schools decreased the gap in years of education by about 1.5 years. The coefficient estimate for females carries the same sign but is much smaller in absolute terms and insignificantly different from zero. Note also that standard errors on estimates are very similar to those reported in tables 5.

## 6.2 Alternative samples

Recall that only one state, Lower Saxony, introduced the reform and that we found some evidence that the effect of parental education on own education was more important in that state (section 4.3). In this subsection, we therefore exclude other confounding factors by carefully altering underlying samples. Table 7 reports results from estimating (4.5), the specification that includes state-background-, cohort-background-, and state-cohort-fixed effects, with years of education as the outcome variable. Baseline estimates from column (3) of table 5 are reproduced in column (1) for reference. Estimates for females and males are reported in panels A and B, respectively. Since we found evidence of an effect on the gradient only for males, we are interested mainly in showing that this finding is robust. We report estimates of the effect for females mainly for completeness.

A potential confounding factor in our setting is pre-primary education. Pre-school enrollment

Table 7: Impact of the reform on inequality in educational outcomes for males. Robustness to alternative sample definitions.

	Baseline estimate (1)	Excl. BW and RP (2)	Excl. BV, BW and RP (3)	Excl. BW, BV, RP, and cohorts born before 1945 (4)	Excl. rural up-bringing (5)	Excl. migrants (6)	Excl. incomplete families (7)	BV, NRW, and SH (8)	Excl. partially treated (9)
<i>Panel A. Females.</i>									
Reform $\times$ ed. parents	-0.332 (0.385)	-0.621 (0.389)	-0.808* (0.413)	-0.221 (0.457)	0.526 (0.550)	-0.305 (0.415)	-0.199 (0.416)	0.474 (0.459)	-0.316 (0.459)
$LS_s \times I[b = high] \times c$	0.046 (0.044)	0.026 (0.044)	-0.002 (0.044)	-0.004 (0.089)	0.071 (0.044)	0.040 (0.047)	0.042 (0.053)	0.083 (0.056)	0.046 (0.044)
Observations	6,630	4,890	3,444	2,430	4,137	5,935	5,640	2,748	4,645
R-squared	0.222	0.237	0.244	0.238	0.247	0.228	0.230	0.223	0.240
<i>Panel B. Males, years of education.</i>									
Reform $\times$ ed. parents	-1.509*** (0.541)	-1.458*** (0.530)	-1.463*** (0.548)	-1.566*** (0.608)	-1.794** (0.707)	-1.561*** (0.525)	-1.647*** (0.622)	-1.629*** (0.631)	-1.625*** (0.647)
$LS_s \times I[b = high] \times c$	0.010* (0.006)	0.007 (0.006)	0.001 (0.006)	0.003 (0.011)	0.018** (0.007)	0.008 (0.006)	0.008 (0.007)	0.014 (0.009)	0.010* (0.006)
Observations	6,180	4,534	3,174	2,245	3,799	5,520	5,285	2,565	4,349
R-squared	0.192	0.194	0.211	0.205	0.216	0.197	0.207	0.219	0.197
<i>Panel B2. Males, university eligibility.</i>									
Reform $\times$ ed. parents	-0.202** (0.086)	-0.186** (0.082)	-0.184** (0.087)	-0.184* (0.093)	-0.218** (0.102)	-0.224** (0.087)	-0.194** (0.095)	-0.244** (0.108)	-0.178* (0.106)
$LS_s \times I[b = high] \times c$	0.010* (0.006)	0.007 (0.006)	0.001 (0.006)	0.003 (0.011)	0.018** (0.007)	0.008 (0.006)	0.008 (0.007)	0.014 (0.009)	0.010* (0.006)
Observations	6,180	4,534	3,174	2,245	3,799	5,520	5,285	2,565	4,349
R-squared	0.137	0.139	0.156	0.147	0.162	0.142	0.148	0.154	0.142
<i>Panel B3. Males, university graduation.</i>									
Reform $\times$ ed. parents	-0.217** (0.094)	-0.203** (0.091)	-0.208** (0.093)	-0.260*** (0.099)	-0.224* (0.129)	-0.223** (0.094)	-0.263*** (0.100)	-0.255** (0.105)	-0.255** (0.106)
$LS_s \times I[b = high] \times c$	0.015* (0.008)	0.014 (0.008)	0.010 (0.007)	-0.011 (0.015)	0.021** (0.009)	0.013 (0.009)	0.014 (0.009)	0.018* (0.010)	0.015* (0.008)
Observations	6,180	4,534	3,174	2,245	3,799	5,520	5,285	2,565	4,349
R-squared	0.137	0.134	0.147	0.164	0.163	0.143	0.149	0.170	0.150

Standard errors clustered at the state-cohort-level in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. All regressions include full sets of dummy variables for the size of respondents' childhood place of residence and migrant status. Based on GSOEP v29-data.



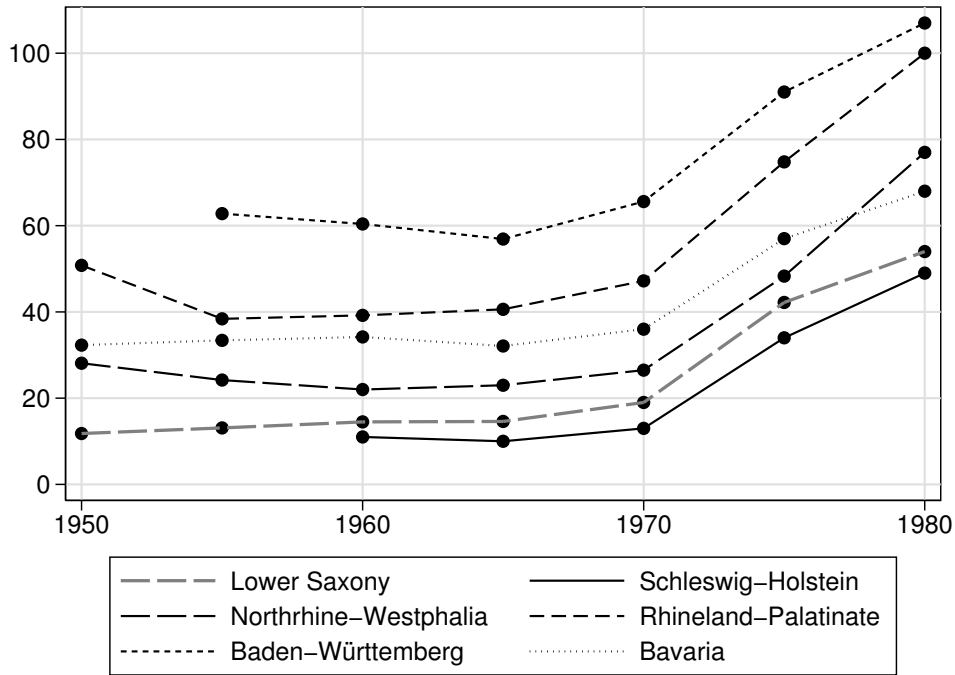


Figure 3: Pre-school capacity per 100 children between the ages of three and six. *Source: Erning et al. (1987, p. 37).*

has repeatedly been shown to affect both cognitive and behavioral outcomes, particularly among disadvantaged children.<sup>28</sup> Hence, it is conceivable that the gradient depends on pre-school enrollment: if educated (uneducated) parents have a higher propensity to enroll their children, we would expect to see a higher (lower) gradient. More recently, [Schütz et al. \(2008\)](#) have argued that the length of pre-primary enrollment is positively associated with a less important effect of parental background and that the effect of the enrollment rate on the gradient is non-linear with an increasing gradient initially and a decreasing gradient at higher levels of enrollment.

Our data do not provide information about whether respondents attended pre-primary education and official records by state are not readily available. A proxy for enrollment, the number of places available per one hundred children between the ages of three and six—what we will refer to as ‘pre-school capacity’—are tabulated, however, in [Erning et al. \(1987, p. 37\)](#). We reproduce these data in figure 3. Note that while there are differences in the level of pre-school capacity, there are no major differences in trends over much of the time period we investigate. Capacity was comparatively low in Lower Saxony, higher in the south and southwest, and increased substantially in all states from 1970 onwards.

<sup>28</sup>[Cunha et al. \(2010\)](#) and [Blau and Currie \(2006\)](#) provide reviews. For Germany, [Schlotter \(2011\)](#) finds some evidence that pre-school education positively affects assertiveness and the ability to make friends, whereas [Schlotter and Wößmann \(2010\)](#) find no evidence for an effect on reading test scores.

This suggests that if the effect of aggregate enrollment on outcomes were linear, trends would not be systematically correlated with the introduction of the reform. Differences in trends may still confound our analysis, however, if the marginal effect of enrollment depends on initial capacity. We therefore ran a supplementary analysis based on matching the data plotted in figure 3 with our microdata (see appendix B). We show that our data are consistent with a u-pattern, not an inverted-u pattern as reported by Schütz et al.<sup>29</sup> Intriguingly, we also find that differences in pre-school capacity across states explain differences in the gradient we encountered in section 4.3.<sup>30</sup>

Given these results, we investigate the sensitivity of our findings to the exclusion of states that differed markedly in terms of pre-school capacity from the sample: in column (2), we report results for a sample from which observations from both Baden-Württemberg and Rhineland-Palatinate, two states that had much higher levels of pre-school capacity (see figure 3), were removed. In column (3), we additionally exclude Bavaria. As we have no information for pre-school capacity prior to 1950, we exclude observations from these three states as well individuals born before 1945 in column (4).<sup>31</sup> Our estimates of the effect of the reform on the gradient remain unaltered throughout. Therefore, we rule systematic differences in pre-school enrollment as an explanation for our findings.

Another factor that is sometimes linked to gradients in educational outcomes is access. While Heineck and Riphahn (2009) find no effect of Germany’s educational expansion on the intergenerational transmission of education for cohorts born between 1929 and 1978 on average, Lower Saxony may be different as it is thinly populated in comparison to other states.<sup>32</sup> For instance, Kramer (2002) documents an increase in access to upper vocational and academic track schools especially in rural areas of Lower Saxony following the construction of new schools during the 1960s and early 1970s—roughly the same time period over which the reform was introduced. Therefore, one may suspect that systematic variation in access to different types of schools confounds our results.

Again, this can be investigated by restricting the underlying sample: if our results were driven by students in rural areas gaining access to upper vocational and academic track schools, we would find no effect of the reform once we exclude these students. We therefore restrict the sample to individuals that grew up in urban and semi-urban areas, roughly two-thirds of the individuals in our sample (column (5) of table 7). While only significant at the ten percent-level, the point estimate hardly changes.

We run further sub-sample robustness checks excluding migrants (column (6)), individuals that were brought up with at least one biological parent absent ((7)), and—in order to show

---

<sup>29</sup>This pattern is consistent with a higher propensity for uneducated parents to have their children attend pre-school educational institutions. It is conceivable that the propensity changed over time, i.e. that uneducated, working class-parents had a higher propensity to have their children attend pre-school in the past.

<sup>30</sup>Since mothers’ time seems more likely to be a substitute for pre-school enrollment, our finding in section 4.3 that pre-reform differences in the gradient are driven by the salience of mothers’ education is also consistent with differences in pre-primary capacity accounting for differences in the gradient.

<sup>31</sup>In Germany, children usually attend “kindergarten” between the ages of three and six.

<sup>32</sup>It is the second-largest state of Germany in terms of area yet only the fourth-largest in terms of inhabitants.

that our results do not hinge on including any particular state as a control—individuals that received schooling in either North Rhine-Westphalia, Bavaria, or Schleswig-Holstein ((8)).<sup>33</sup> All these alterations do not affect our results.

A final test is for self-selection into OS schools. The introduction of OS schools in Lower Saxony took almost one decade and during this time period selection may have played a role. Given our findings above, it seems likely that educated parents would have decided to send their children to schools in the old, three-tiered system as long as this was possible and one would therefore expect self-selection to attenuate the effect of the reform on the gradient. In any case, the co-existence of two tracking regimes potentially produces outcomes that differ from those under complete de-tracking.<sup>34</sup> We therefore re-estimate the model using only observations from cohorts that were either fully treated or fully untreated, that is, cohorts born before 1962 or after 1971. Compared to the baseline, the resulting estimate in column (8) suggests a slightly more pronounced, negative effect of the reform on the gradient.

Overall, our findings are remarkably robust to changes in the underlying sample. We also rule out differences in levels and trends in pre-primary enrollment and the concomitant expansion of access to educational institutions as explanations for our findings.

### 6.3 Placebo test

We now present results from a placebo test that allows us to rule out several alternative explanations for our findings above. The idea is to test whether the reform was associated with a decrease in the gap in an alternative indicator of human capital that cannot plausibly have been the consequence of the reform. One such indicator that received considerable attention in the literature recently is human body height. Taller people earn more on average, they do better on cognitive tests, and they tend to live longer (Deaton, 2007). Body height also predicts IQ and self-esteem and is negatively correlated with depression symptoms and behavioral problems (von Hinke Kessler Scholder et al., 2013). In our data, we find that conditional on our usual set of controls as well as age, an increase in height by ten centimeters is associated with an increase by about three-fifths of a year in education. This effect is very similar across gender (results not shown). We further find that having educated parents is associated with a one and 1 1/2 centimeters increase in body height for females and males, respectively. The effect does not vary significantly between states prior to the reform.

Body height is determined to a large degree before the age of ten<sup>35</sup>, the age at which the reform we study may have interfered with physiological development, and is a consequence of (net) nutrition and genetic potential. It is not plausible that an education reform that affected tracking at the age of ten could have had an impact on human body height. If, however, we were to find that the effect of parental education on the gradient in height changed with the

---

<sup>33</sup>In other words, we keep as controls only the states that we are excluding in column (2).

<sup>34</sup>Recall that this is why Hesse was excluded, a state in which the two systems exist next to each other.

<sup>35</sup>While puberty is a significant event in human growth, the growth spurts associated with this period accounts only for 17–18 percent of final adult body height (Abbassi, 2015).

Table 8: Placebo test: POLS and BE estimates of the impact of the reform on gradient between parental education and body height.

	Females		Males	
	POLS (1)	BE (2)	POLS (3)	BE (4)
Reform $\times$ ed. parents	-0.485 (0.565)	-0.476 (1.016)	0.507 (0.670)	0.461 (1.196)
Observations	21,421	21,421	20,005	20,005
R-squared	0.13	0.14	0.14	0.15

Asymptotic standard errors in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. All regressions include state-background-, state-cohort-, cohort-background-, and age-fixed effects. Further controls are dummies capturing the size of respondents' childhood place of residence and migrant status. Based on GSOEP v29-data.

introduction of de-tracking, it would seem that unobserved factors are driving our results. In particular, such as finding would seem to indicate unobserved changes in the composition of the groups that we are comparing. The present placebo test allows us to rule out this possibility.

The empirical specification we use for the placebo test is similar to those used above. However, body height is only measured infrequently in our dataset and varies with age. We therefore use the panel-dimension of the dataset and employ pooled OLS (POLS)- and between effects (BE)-estimators. We control for age by including a full set of fixed effects. Other than that, the specification is the same as (4.2), that is, we use our standard set of controls in addition to state-background-, state-cohort-, and cohort-background-fixed effects. We restrict the sample to respondents that are at least 21 years of age and born between 1940 and 1984. Standard errors are cluster around the individual-identifier for POLS estimates. Results are reported in table 8 for females and males.<sup>36</sup>

Overall, there is no indication that the reform is associated with a change in the gradient between parental education and body height. The coefficient on the interaction term is positive for males (columns (3) and (4)), indicating that the gap in terms of body height may actually have widened. However, the coefficient is not significantly different from zero at conventional levels of statistical significance. We conclude that there are no major changes in environmental conditions or even genetic background variables between cohorts that differed systematically between Lower Saxony and other West-German states.

<sup>36</sup>As is appropriate in the case of a placebo test, we use the *least* conservative options with regard to standard errors and estimation methods. In particular, we use weighted least squares when we estimate the BE equation and asymptotic standard errors (rather than standard errors clustered around the state-cohort- or—in the case of POLS—individual-identifier).

## 7 Discussion

Results presented in previous sections suggest that the reform achieved for male students what it was intended to achieve: it attenuated the importance of parental background for students' educational prospects. There is no evidence for a trade-off between efficiency and equity; the reform did not seem to have lowered overall educational attainment.

The effect on the gradient in total years of education for males is significant in economic terms given that tracking was delayed by only two years: our point estimates suggest that the gradient decreased by between 1.2 and 1.5 years, at least one-half of the (unconditional) gradient that we observe in Lower Saxony prior to the reform.

To further put these results into perspective, we can compare them to findings for wage effects of the Swedish reform studied in [Meghir and Palme \(2005\)](#). Returns to one additional year of education for males in Germany are usually estimated to be around seven percent ([Lauer and Steiner, 2000](#); [Ammermüller and Weber, 2005](#)). The implied changes in hourly wages for males with educated and uneducated parents that we would expect from changes in years of education (based on results in column (1) of table 5) are  $(0.442 - 1.031) \times 0.07 \approx -4.1$  percent for males with educated parents and  $0.442 \times 0.07 \approx 3.1$  percent for males with uneducated parents. While these effects would neither be very large nor negligible, they are well in line with findings in [Meghir and Palme \(2005\)](#) who report estimates of  $-7.7$  percent and  $3.1$  percent, respectively.<sup>37</sup> An important difference is that we find no effects for females while [Meghir and Palme](#) find a positive effect for females with uneducated fathers (about 3.8 percent) and a negative effect for females with educated fathers (about  $-4.2$  percent).

While the effect on the gradient we find for boys seems large, we pointed out several reasons why it is likely a lower bound estimate: first, our matching of individuals to states in which they received schooling will be erroneous to some degree and this would tend to attenuate our coefficient estimate of interest (see section 4 and, in particular, appendix A). Second, we find that the gradient may have increased slightly in Lower Saxony prior to the reform relative to other states (section 4.3.1). If this would have continued in the absence of the reform, we would tend to underestimate the negative effect on the gradient. Finally, we included for the most part cohorts that were on partially affected by the reform. If we assume that parents were in a position to make a strategic choice on behalf of their children, keeping these cohorts in the estimation sample will likely tend to bias our estimate towards zero.

Why do we find an effect only for males? A plausible explanation would be that tracking age interacts with gender differences in the development of non-cognitive skills. This would require that educated parents have an advantage in the production of these skills, that the resulting advantage dissipates over time (for instance, through in-class peer effects), and that at least one of the above is more relevant for boys than for girls. If these skills enhance learning or educators reward these skills, we would expect to find that delayed tracking is associated with a

---

<sup>37</sup>Remember, however, that the Scandinavian did also encompass transfer payments and an increase in compulsory schooling.

less prominent role of parental education for boys.

Differences in the levels and trajectories of cognitive and non-cognitive development between boys and girls are well-established in the literature. [Matthews et al. \(2009\)](#) find evidence for gender differences in self-regulation, the ability to control behavior, cognitions, and emotions, for kindergarten children with females outperforming boys in self-control but not in academic achievement. [Lenroot et al. \(2007\)](#) find important differences in the trajectories of brain development between boys and girls with boys generally trailing behind.

Recently, gender differences in non-cognitive skills have directly been linked to the gender gap in academic achievement. Based on data from the US, [DiPrete and Jennings \(2012\)](#) show that while there are no gender gaps in the returns to social and behavioral skills for children from kindergarten through fifth grade, girls lead boys by nearly 0.4 standard deviations at the start of kindergarten. They demonstrate that this gap grows over time and explains a considerable fraction of the gender gap in academic outcomes at that age. Moreover, these skills are significantly related to parents' socio-economic status and the presence of a father. While they reckon that this may reflect teachers rewarding social and behavioral skills, they also find evidence that these skills enhance learning. This is in line with [Kerr et al. \(2013\)](#) who find benefits from comprehensive schooling in terms of higher cognitive skills for students from disadvantaged backgrounds.

## 8 Conclusion

Recent research suggests that design features of education systems are an important determinant of the intergenerational transmission of education, the effect of parental on own education. While tracking has often been found to increase variation in outcomes and the salience of parental background, the practice is also often credited with increasing the efficiency of teaching through an increase in class homogeneity.

In this study, we investigate whether and how the introduction of delayed tracking in one of Germany's federal states in the 1970s affected educational outcomes of individuals differentiated by gender and parental education. Based on a difference-in-differences estimation strategy, we find no evidence that the reform led to a decrease in educational attainment for cohorts affected, that is, there is no evidence that tracking increases efficiency. We present strong evidence, however, that the reform is associated with increased equality of opportunity. The effect is entirely driven by males: the reform benefited males with uneducated parents at the expense of males with educated parents. These findings are robust to a large number of alternative specifications and restrictions imposed on the underlying sample. There is no indication that the reform had any effect on females.

We find no evidence that the reform affected parental involvement, one of its main aims. It seems plausible that the gender differences we observe are related to changes in the development of non-cognitive skills for males. However, more work on the precise transmission channel is

clearly warranted. The interaction between gender differences in skill development and tracking seems an interesting alley for future research.

## References

- Abbassi, V. (2015). Growth and Normal Puberty. *Pediatrics*, 102:507–511.
- Ammermüller, A. and Weber, A. M. (2005). Educational Attainment and Returns to Education in Germany—An Analysis by Subject of Degree, Gender and Region. *Centre for European Economic Research (ZEW) Discussion Paper 05-17*.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, New Jersey.
- Avenarius, H., Döbert, H., Knauss, G., Weishaupt, H., and Weiß, M. (2001). Stand und Perspektiven der Orientierungsstufe in Niedersachsen. Technical report, Deutsches Institut für Internationale Pädagogische Forschung.
- Basu, K. and Foster, J. E. (1998). On Effective Literacy. *Economic Journal*, 108(451):1733–1749.
- Bauer, P. and Riphahn, R. T. (2006). Timing of School Tracking as a Determinant of Intergenerational Transmission of Education. *Economics Letters*, 91:90–97.
- Baumert, J., Trautwein, U., and Artelt, C. (2003). Schulumwelten—Institutionelle Bedingungen des Lehrens und Lernens. In Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., and Weiß, M., editors, *PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*, pages 261–331. Leske + Budrich, Opladen, Germany.
- Bedard, K. and Cho, I. (2010). Early Gender Test Score Gaps across OECD Countries. *Economics of Education Review*, 29:348–363.
- Betts, J. R. (2010). The Economics of Tracking in Education. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 3, pages 341–381. North Holland, Amsterdam.
- Betts, J. R. and Shkolnik, J. R. (2000). Key Difficulties in Identifying the Effects of Ability Grouping on Student Achievement. *Economics of Education Review*, 19(1):21–26.
- Blau, D. and Currie, J. (2006). Pre-School, Day Care, and After-School Care: Who's Minding the Kids? In Hanushek, E. A. and Welch, F., editors, *Handbook of the Economics of Education*, volume 2, pages 1163–1278. Elsevier, Amsterdam.
- Brunello, G. and Checchi, D. (2007). Does School Tracking Affect Equality of Opportunity? New International Evidence. *Economic Policy*, 22(52):781–861.
- Brunello, G., Giannini, M., and Ariga, K. (2007). The Optimal Timing of School Tracking: A General Model with Calibration for Germany. In Woessmann, L. and Peterson, P., editors, *Schools and the Equal Opportunity Problem*, pages 129–156. The MIT Press, Cambridge, Mass.



- Bundesministerium der Finanzen (2015). The Federal Financial Equalisation System in Germany. [https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Oeffentliche\\_Finzen/Foederale\\_Finanzbeziehungen/Laenderfinanzausgleich/Eng-Der-Bundesstaatliche-FAG.pdf?\\_\\_blob=publicationFile&v=1](https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Oeffentliche_Finzen/Foederale_Finanzbeziehungen/Laenderfinanzausgleich/Eng-Der-Bundesstaatliche-FAG.pdf?__blob=publicationFile&v=1). Accessed: 09/28/2015.
- Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2010). Interpreting the Evidence on Life Cycle Skill Formation. In Hanushek, E. A. and Welch, F., editors, *Handbook of the Economics of Education*, volume 1, pages 697–812. North Holland, Amsterdam.
- Deaton, A. (2007). Height, Health, and Development. *Proceedings of the National Academy of Sciences*, 104(33):13232–13237.
- Deutscher Ausschuss (1959). Empfehlungen zur Umgestaltung und Vereinheitlichung des Allgemeinbildenden Öffentlichen Schulwesens. In Bohnekamp, H., Dirks, W., and Knab, D., editors, *Empfehlungen und Gutachten des Deutschen Ausschusses für das Erziehungs- und Bildungswesen 1953–1965*, pages 59–115. Stuttgart, Germany.
- Deutscher Bildungsrat (1970). Empfehlungen der Bildungskommission: Strukturplan für das Bildungswesen. Technical report, Deutscher Bildungsrat, Stuttgart.
- Ding, W. and Lehrer, S. (2007). Do Peers Affect Student Achievement in China’s Secondary Schools. *Review of Economics and Statistics*, 89(2):300–312.
- DiPrete, T. A. and Jennings, J. (2012). Social/Behavioral Skills and the Gender Gap in Educational Achievement. *Social Science Review*, 41:1–15.
- Duflo, E., Dupas, P., and Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking. *American Economic Review*, 101:1739–1774.
- Dustmann, C., Puhani, P. A., and Schönberg, U. (2014). The Long-Term Effects of Early Track Choice. *IZA Discussion Paper No. 7897*.
- Epple, D. and Romano, R. (2011). Peer Effects in Education: A Survey of the Theory and Evidence. In Benhabib, J., Jackson, M. O., and Bisin, A., editors, *Handbook of Social Economics*, volume 1B, pages 1053–1163. North Holland, Amsterdam.
- Erning, G., Neumann, K., and Reyer, J. (1987). *Geschichte des Kindergartens: Institutionelle Aspekte, Systematische Perspektiven, Entwicklungsverläufe*, volume 2. Lambertus, Freiburg, Germany.
- Haenisch, H. and Ziegenspeck, J. W. (1977). *Die Orientierungsstufe*. Beltz Verlag, Weinheim, Germany, and Basel, Switzerland.
- Hall, C. (2012). The Effects of Reducing Tracking in Upper Secondary School: Evidence from a Large-Scale Pilo Scheme. *Journal of Human Resources*, 47:237–269.

- Hanushek, E. A. and Woessmann, L. (2006). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries. *The Economic Journal*, 116(510):C63–C76.
- Heineck, G. and Riphahn, R. (2009). Intergenerational Transmission of Educational Attainment in Germany—The Last Five Decades. *Journal of Economics and Statistics (Jahrbücher fuer Nationalökonomie und Statistik)*, 229(1):36–60.
- Hornich, K. (1976). *Praxis der Orientierungsstufe*. Verlag Herder.
- Jürges, H. and Schneider, K. (2011). Why Young Boys Stumble: Early Tracking, Age and Gender Bias in the German School System. *German Economic Review*, 12(4):371–394.
- Kerr, S. P., Pekkarinen, T., and Uusitalo, R. (2013). School Tracking and Development of Cognitive Skills. *Journal of Labor Economics*, 31(3):577–602.
- Kramer, W. (2002). *50 Jahre Schulenwicklung in Niedersachsen*. Bibliotheks- und Informationssystem der Universität Oldenburg, Oldenburg, Germany.
- Landesamt für Statistik Niedersachsen (2014). Amtliche Schulstatistik. Various years.
- Lauer, C. and Steiner, V. (2000). Returns to Education in West Germany - An Empirical Assessment. *ZWE Discussion Papers No. 00-04*.
- Lavy, V., Paserman, D., and Schlosser, A. (2011). Inside the Black Box of Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *The Economic Journal*, 122(559):208–237.
- Lehmann, R. and Peek, R. (1997). Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Jahrgangsstufe an Hamburger Schulen. Bericht über die Untersuchung im September 1996.
- Lenroot, R. K., Gogrtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., Blumenthal, J. D., Lerch, J., Zijdenbos, A. P., Evans, A. C., Thompson, P. M., and Giedd, J. N. (2007). Sexual Dimorphism of Brain Developmental Trajectories During Childhood and Adolescence. *NeuroImage*, 36:1065–1073.
- Lyle, D. S. (2009). The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point. *American Economic Journal: Applied Economics*, 1(4):69–84.
- Malamud, O. and Pop-Eleches, C. (2011). School Tracking and Access to Higher Education Among Disadvantaged Groups. *Journal of Public Economics*, 95:1538–1549.
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60:531–542.

- Matthews, J., Ponitz, C. C., and Morrison, F. J. (2009). Early Gender Differences in Self-Regulation and Academic Achievement. *Journal of Educational Psychology*, 101(3):689–704.
- Meghir, C. and Palme, M. (2005). Educational Reform, Ability, and Family Background. *American Economic Review*, 95(1):414–424.
- Mühlenweg, A. M. (2008). Educational Effects of Alternative Secondary School Tracking Regimes in Germany. *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 128(3):351–379.
- Mühlenweg, A. M. and Puhani, P. A. (2010). Persistence of the School Entry Age Effect in a System of Flexible Tracking. *Journal of Human Resources*, 45:407–435.
- Pekkarinen, T., Uusitalo, R., and Kerr, S. (2009). School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform. *Journal of Public Economics*, 93(7–8):965–973.
- Piopiunik, M. (2013). The Effects of Early Tracking on Student Performance: Evidence from a School Reform in Bavaria. *Ifo Working Paper No. 153*.
- Pischke, J.-S. and Manning, A. (2006). Comprehensive Versus Selective Schooling in England and Wales: What Do We Know? *NBER Working Paper No. 12176*.
- Puhani, P. A. and Weber, A. M. (2007). Does the Early Bird Catch the Worm? Instrumental Variables Estimates of the Educational Effects of Age at School Entry in Germany. *Empirical Economics*, 32:359–386.
- Rösner, E. (1981). *Schulpolitik durch Volksbegehren. Analyse eines gescheiterten Reformversuchs*. Beltz, Weinheim, Germany.
- Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116(2):681–704.
- Schlotter, M. (2011). Age at Preschool Entrance and Noncognitive Skills before School—An Instrumental Variable Approach. *Ifo Institute for Economic Research Working Paper No. 112*.
- Schlotter, M. and Wößmann, L. (2010). Frühkindliche Bildung und Spätere Nicht-kognitive Fähigkeiten: Deutsche und Internationale Evidenz. *Ifo Institute for Economic Research Working Paper No. 91*.
- Schneeweis, N. and Zweimüller, M. (2014). Early Tracking and the Misfortune of Being Young. *Scandinavian Journal of Economics*, 116(2):394–428.
- Schnepf, S. V. (2002). A Sorting that Fails? The Transition from Primary to Secondary School in Germany. *UNICEF Innocenti Working Paper No. 92*.

- Schuchart, C. (2006). *Orientierungsstufe und Bildungschancen: Eine Evaluationsstudie*. Waxmann, Münster, Germany.
- Schütz, G., Ursprung, H. W., and Wößmann, L. (2008). Education Policy and Equality of Opportunities. *Kyklos*, 61(2):279–308.
- Segal, C. (2008). Classroom Behavior. *Journal of Human Resources*, 43(4):783–814.
- von Friedeburg, L. (1992). *Bildungsreform in Deutschland: Geschichte und Gesellschaftlicher Widerspruch*. Suhrkamp Verlag.
- von Hinke Kessler Scholder, S., Smith, G. D., Lawlor, D. A., Propper, C., and Windmeijer, F. (2013). Child Height, Health and Human Capital: Evidence Using Genetic Markers. *European Economic Review*, 57:1–22.
- Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements. *Schmollers Jahrbuch*, 127(1):139–169.
- Waldinger, F. (2007). Does Ability Tracking Exacerbate the Role of Family Background for Students' Test Scores?
- Woessmann, L. (2010). Institutional Determinants of School Efficiency and Equity. *Jahrbücher für Nationalökonomie und Statistik*, 230(2):234–270.
- Ziegenspeck, J. W. (2000). *Handbuch Orientierungsstufe: Sachstandsbericht und Zwischenbilanz*. Klinkhardt, Bad Heilbrunn, Germany.

## A Inferring the state of schooling at age ten

In narrowing down the state in which individuals received their schooling at the age of ten, we rely both on additional items in the dataset as well as on the time dimension of the dataset. We proceed as follows:

1. We impute the state of the last school visit whenever the exact state cannot be inferred. This information was only gathered in one year, so there are many individuals in the dataset for which it is not available.
2. If neither the exact state nor the state of the last school visit is available, we resort to information from a question about the childhood state of residence. All individuals were asked whether they still live in the same place in which they lived during their childhood. While ‘childhood’ does not seem particularly well defined, we impute the first state in which an individual was encountered whenever the answer is ‘yes.’
3. If all of the above fails, we impute the current state.
4. Finally, we remove individuals from our sample that match one of the following criteria: first, all individuals were asked about the place in which they lived in 1989, shortly before Germany’s re-unification. If somebody states that she lived in East Germany and was born before 1978, it is nearly impossible that she received her schooling in one of the Western German states. We thus drop all these observations from our sample. Second, the dataset contains information about the year in which an individual migrated to Germany. Together with information on birth years, we can thus calculate the age at which an individual migrated. We drop all migrants that moved to Germany only after they turned eleven.

Table 9 reports the frequencies with which individuals were classified according to the above steps by treatment status and gender. It shows that for the majority of individuals, we make the assumption that the state of last school visit is also the state they lived in at the age of ten. Less than five percent are actually observed in the data once at the age of ten and then, again, when they are at least 28 years old. Around 15 percent are classified based on the observation that they are still living in the same state in which they have also lived during their childhood. Only around eleven to 12 percent of individuals are classified based on information about the state in which they have first been observed in the data.

## B Differences in pre-primary enrollment rates and the intergenerational transmission of education

Recent studies have found a non-linear effect of pre-primary education on equality of opportunity in education (Schütz et al., 2008). While there do not seem to be any significant differences in trends (see figure 3), there are large differences in enrollment rates between states around the

Table 9: Provenance of information on state of school visit at the age of ten.

	Other states			Lower Saxony		
	All	Females	Males	All	Females	Males
...state at age ten.	0.02	0.02	0.02	0.02	0.02	0.01
...state of last school visit.	0.38	0.38	0.39	0.41	0.42	0.40
...childhood state.	0.34	0.34	0.34	0.31	0.31	0.30
...first state in which observed.	0.26	0.26	0.25	0.26	0.25	0.28
Observations	98,324	50,662	47,662	16,520	8,486	8,034

Statecodes as in notes to figure 1. Based on GSOEP v29-data.

time the reform was implemented. Therefore, changes in enrollment rates over time may still confound our results. In this appendix, we will analyze the link between enrollment rates in pre-primary and equality of opportunity in our data. We show that differences in rates account for differences in the gradient both prior to the reform (as reported in section 4.3) and that (2), in contrast to findings reported in Schütz et al. (2008), the relationship between the gradient and pre-primary enrollment follows a u-shape. It therefore becomes important to account for these non-linearities, which we do in section 6.

We proceed as follows: first, we test whether pre-primary enrollment rates (proxied by capacity per child as depicted in figure 3) have an effect on educational outcomes and the gradient between parental and own education. We focus on the subsample of males and regress outcomes on an indicator of having educated parents interacted with state dummies, i.e. we estimate regressions similar to those reported in section 4.3. However, we also include state- and cohort-fixed effects as well as controls for childhood place of residence and migrant status. The idea is to show that differences in the gradient can be explained by differences in pre-primary enrollment rates. Therefore, we include in a second step data on pre-primary enrollment. Since data at the state-cohort levels is only available at quinquennial intervals, we always restrict the sample to individuals pertaining to cohorts for which the pre-primary enrollment rate is available at the time they were five to six years old.<sup>38</sup>

Results are reported in table 10. In column (1), we focus on the subsample of individuals that turned ten before the introduction of the reform. Results are very similar to those reported in table 3: the conditional gradient for Lower Saxony, the reference category, is about 3.1 years and highly significant. There is again some evidence for differences in the pre-reform gradient across states. Although we cannot reject that coefficients on interaction terms are jointly zero ( $p$ -value of 0.282), estimates for some states are negative and individually significant. This suggests that parental education was less important in these states compared to Lower Saxony.

Next, we include in addition the pre-primary enrollment rate, its square, and interactions of

<sup>38</sup>To give an example, we have data on the pre-primary enrollment rate in 1960 for most states. Hence, we include the cohorts born in late 1955 or early 1956. Note that this dramatically reduces the total number of observations available.

Table 10: Effect of pre-primary enrollment on outcomes and the gradient between parental and own education for males.

	Pre-reform			All but Lower Saxony post-reform		
	(1)	(2)	(3)	(4)	(5)	(6)
Educated parents	2.420*** (0.440)	3.909 (4.913)	3.605*** (1.102)	2.356*** (0.431)	3.496*** (0.597)	3.324*** (0.622)
Educated parents × BV	-0.985 (0.754)	0.898 (4.318)		-0.778 (0.557)	0.385 (0.753)	
Educated parents × BW	-0.935 (0.756)	3.256 (7.403)		-0.209 (0.602)	0.621 (0.829)	
Educated parents × NRW	-0.669 (0.667)	0.305 (2.610)		-0.785 (0.522)	0.022 (0.697)	
Educated parents × RP	-1.112 (0.676)	1.545 (5.204)		-0.927 (0.747)	0.048 (0.851)	
Educated parents × SH	0.293 (0.857)	0.005 (1.314)		-0.203 (1.010)	0.087 (0.991)	
Pre-primary enrollment		-2.459 (8.570)	-2.379 (12.574)		11.001*** (3.569)	10.026*** (3.529)
Pre-primary enrollment squared		9.154 (12.222)	7.302 (16.476)		-5.507*** (1.245)	-5.081*** (1.276)
Educated parents × enrollment		-11.167 (41.050)	-10.362 (7.623)		-9.690*** (3.099)	-8.407*** (2.379)
Educated parents × enrollment squared		2.715 (47.166)	11.365 (11.018)		9.134*** (2.579)	8.508*** (2.114)
<i>Results of F-tests for joint significance (p-values).</i>						
Education-state-interactions	0.420	0.993		0.565	0.876	
Education-enrollment-interactions		0.587	0.269		0.000	0.000
Observations	564	564	564	931	931	931
R-squared	0.157	0.159	0.158	0.157	0.167	0.166

Standard errors clustered at the state-cohort-level in parentheses. \*, \*\*, and \*\*\* denote significance at the ten-, five-, and one-percent level, respectively. All regressions include sets of dummies capturing the size of respondents' childhood place of residence and migrant status as well as dummies for state of residence at age ten and birth cohort. Based on GSOEP v29-data.

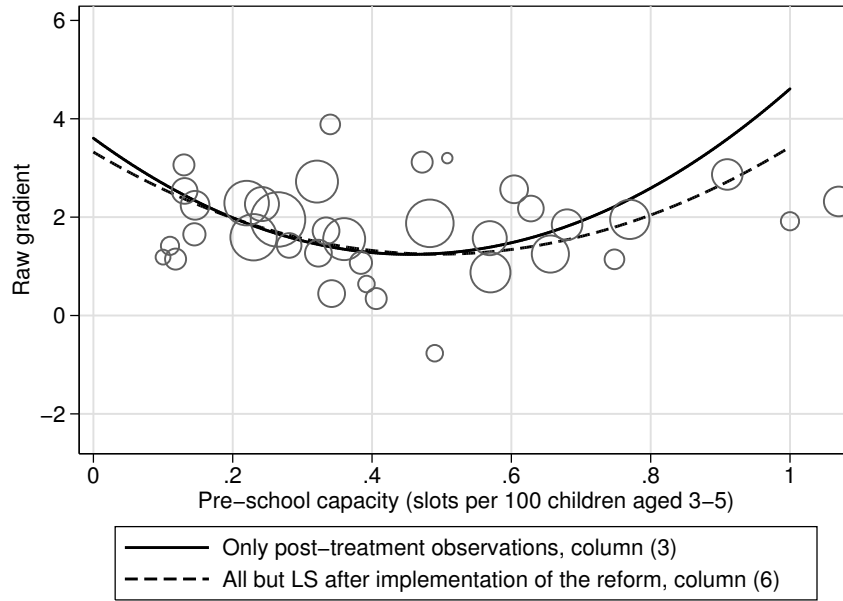


Figure 4: Illustration of the effect of pre-primary enrollment rates on the gradient between parental and own education. Based on results in columns (3) and (6) of table 10. Hollow circles depict state-cohorts for which data on pre-school capacity is available. Only cohorts that went to school in Lower Saxony after the implementation of the reform are excluded. The size of the circles is proportional to the number of observations in the respective state-cohort.

these two variables with our indicator of parental education. Comparing estimates and significance levels between columns (1) and (2), it is apparent that the inclusion of these variables accounts for a large portion of the variation in gradients prior to the reform: all estimates on state-parental education-interactions are now statistically insignificant and most are closer to zero. Thus, it seems that difference in enrollment rates account for differences in gradients prior to the reform. However, we find that we cannot reject that variables capturing the relationship between outcomes, pre-primary enrollment, and parental education are jointly insignificant.

Ignoring column (3) for now, we expand the sample by including cohorts that would have been affected by the reform had they lived in Lower Saxony. Cohorts actually affected are excluded on the basis that our analysis shows that the relationship between parental education and own education changed as a consequence of the reform. Re-running the same regressions, we see that variables capturing pre-primary enrollment still account for differences in the gradient. We also find that all variables that capture effects of pre-primary enrollment are now individually and jointly significant.

Having shown that differences in pre-primary enrollment rates account for differences in gradients across states, we further explore the relationship between rates and the gradient between parental and own education. To do so, we first re-run regressions in columns (2) and (4) but



excluding the state-parental education-interactions. Results are reported in columns (3) and (6) for the two samples, respectively. To illustrate what these coefficient estimates imply, we provide some graphical evidence. Figure 4 plots raw gradients in years of education for all male cohorts and states for which data on pre-primary enrollment is available. The size of the markers indicate the number of observations for each state-cohort. We overlay this plot with predicted gradients from regressions reported in columns (3) and (6) of table 10.

This indicates that the relationship between the gradient and the pre-primary enrollment follows a non-linear, u-shaped curve. An increase in the enrollment rate will initially attenuate differences in outcomes between males with educated parents and those with uneducated parents. However, the relationship turns at an enrollment rate of about 50–60 percent. Further increases are associated with an increase in enrollment rates. Recall that Schütz et al. (2008) find an inverse-u-shaped pattern for their sample of younger German cohorts.

This is not unreasonable if one assumes that pre-primary enrollment is beneficial: cohorts that had children born during the 1950s to 1970s in Germany lived through an episode of near-full employment while lower-class parents today are more likely to be unemployed or to forego enrolling their children for other reasons. Hence, it may be that working class-parents back then had a higher propensity to enroll their children in pre-primary education while, today, they are less likely to enroll their children.

This finding has important implications for our empirical strategy as it suggests that an increase in enrollment rates in states with high levels of enrollment (e.g. Baden-Württemberg) positively affects the gradient while it negatively affects the gradient in states with low levels of enrollment (e.g. Lower Saxony). This is addressed in section 6, where we show that excluding states that differed significantly from Lower Saxony in terms of pre-primary enrollment increases the estimated effect of the reform on the gradient. However, the effect remains large and significant.