

von Werder, Marten; Lange, Simon

Working Paper

Tracking and the intergenerational transmission of education: Evidence from a natural experiment

SOEPPapers on Multidisciplinary Panel Data Research, No. 880

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: von Werder, Marten; Lange, Simon (2016) : Tracking and the intergenerational transmission of education: Evidence from a natural experiment, SOEPPapers on Multidisciplinary Panel Data Research, No. 880, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<http://hdl.handle.net/10419/148552>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

SOEPpapers
on Multidisciplinary Panel Data Research

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

880-2016

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

Simon Lange and Marten von Werder

SOEPPapers on Multidisciplinary Panel Data Research at DIW Berlin

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

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

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

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

Editors:

Jan **Goebel** (Spatial Economics)
Martin **Kroh** (Political Science, Survey Methodology)
Carsten **Schröder** (Public Economics)
Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics, DIW Research Fellow)
Denis **Gerstorff** (Psychology, DIW Research Director)
Elke **Holst** (Gender Studies, DIW Research Director)
Frauke **Kreuter** (Survey Methodology, DIW Research Fellow)
Frieder R. **Lang** (Psychology, DIW Research Fellow)
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Fellow)
Thomas **Siedler** (Empirical Economics, DIW Research Fellow)
C. Katharina **Spieß** (Education and Family Economics)
Gert G. **Wagner** (Social Sciences)

ISSN: 1864-6689 (online)

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

Contact: soeppapers@diw.de



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

Simon Lange[†] Marten von Werder[‡]

December 7, 2016

Abstract

Proponents of tracking argue that the creation of more homogeneous classes increases efficiency while opponents point out 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 effect is driven entirely by changes in the gradient for males and to a large extent by an effect on the likelihood to complete the academic secondary track.

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

*The authors would like to thank Rafael Aigner, Mikkel Barlund, Ronny Freier, Stephan Klasen, Steffen Lohmann, Richard Mansfield, Malte Reimers, Ramona Rischke, Viktor Steiner, and Sebastian Vollmer for valuable comments on earlier versions of this paper. The paper has also benefited from comments by seminar participants in Berlin, Göttingen, and Williamsburg, as well as by participants at the Spring Meeting of Young Economists 2015 in Ghent, Belgium, the Sixth ECINEQ Meeting in Luxembourg, and the 2015 annual conference of the HDCA in Washington, DC.

[†]Corresponding author. The World Bank, Washington, DC. E-mail: slange@worldbank.org.

[‡]Economics Department, Free University of Berlin.

(Betts, 2010). Countries differ widely in the way they track students¹ 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 of 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 between 1972 and 1982 through the introduction of a new intermediate school, the *orientation stage*² (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.³ 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 one to 1.3 years, accounted for by an increase for males with uneducated parents by 0.3 years and a decrease for males with educated parents by about one year. 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 but that the former effect is much larger.

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, differences in the pace and timing of Germany's educational expansion, or changes in family characteristics. Overall, our finding of a negative effect of the reform on the intergenerational transmission of education is remarkably robust to a number of 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

¹For 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).

²In German: *Orientierungsstufe*.

³We define the gradient as the gap in educational outcomes between individuals with educated parents and those with uneducated parents.

socio-economic and institutional factors.⁴ 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.⁵ 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. While their results thus stand in stark contrast to our findings, the setting also differs in many important ways from the one in West Germany.

Recently, a number of studies that analyze country-wide school reforms in Sweden and Finland have also relied on variation across cohorts and space to analyze de-tracking reforms ([Meghir and Palme, 2005](#); [Pekkarinen et al., 2009](#); [Hall, 2012](#); [Kerr et al., 2013](#)). These studies are comparable in terms of the identification strategy used. They also tend to find a positive effect on equality of opportunity. However, there are also important differences: reforms in Sweden affected students only at older ages⁶ and reforms in both Sweden and Finland involved significant

⁴[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 \(2010\)](#).

⁵See [Betts \(2010\)](#) for a review.

⁶A reform in the early 1960s (studied by [Meghir and Palme, 2005](#)) replaced a system of tracking after sixth grade with one in which students would attend a nine-year comprehensiveness school. A subsequent reform

changes to educational institutions that cannot be separated empirically from the effects of de-tracking.⁷ In comparison, changes that accompanied de-tracking in Lower Saxony were arguably less substantial.

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 suggest hypotheses for the gender differences in effects 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:⁸ 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.

Tracks differ in several respects in terms of the quality and quantity of inputs (Dustmann et al., 2016): first, teachers in the academic track usually receive higher wages. Second, their

(studied by Hall, 2012) was aimed at making upper secondary education more comprehensive. This was to be accomplished through an increase in both the academic content and the length of its vocational track.

⁷The first Swedish reform, for instance, entailed an increase in compulsory years of schooling from seven to nine years as well as cash transfers to compensate families for foregone earnings (Meghir and Palme, 2005). The reform in Finland coincided with the abolition of a vast network of private schools which were placed under municipal authority.

⁸Exceptions 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.

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 academic and 21.25 students per teacher in the vocational tracks.

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. \(2016\)](#) 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., 2016](#)) 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.⁹

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 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

⁹Alternatively, 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.

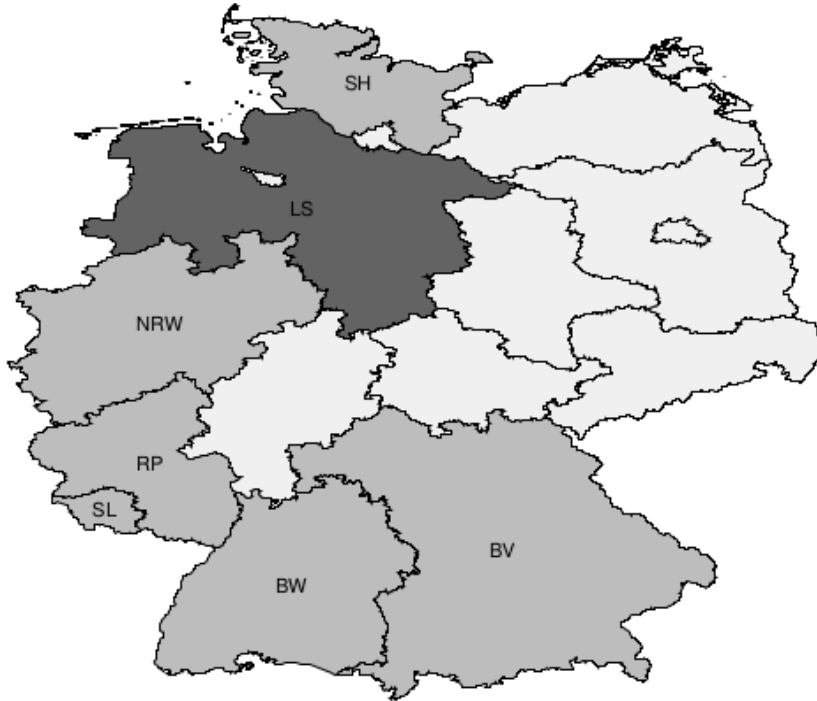


Figure 1: Treatment (dark gray) and control states (light gray). From North to South: SH (Schleswig-Holstein), LS (Lower Saxony), NRW (Northrhine-Westphalia), RP (Rhineland-Palatinate), SL (Saarland), BW (Baden-Württemberg), and BV (Bavaria).

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.¹⁰

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

¹⁰Bremen 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).

¹¹Out of the six remaining states that will serve as controls in our analysis below, two, Northrhine-Westphalia

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 comparisons in our analysis below are between tracking across schools and a system with a strong element of tracking within schools. Third, 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).

Fourth, OS schools employed teachers from all secondary school tracks on a part-time basis and, ideally, each class was to be taught by teachers from all secondary school tracks in at least one subject. This suggests that after the reform, students were exposed to teachers with profound knowledge of the everyday demands at different types of secondary schools prior to the recommendation. The teachers that would now be responsible for the tracking recommendation were also significantly more likely to be male.¹² As we will argue below, this may be one explanation for the pronounced gendered pattern in effects we find.

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 may have been particularly important only from 1980 onward when teachers' recommendations in Lower Saxony lost their binding character. 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.). Hence, if one accepts that teachers have only

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.

¹²Data for the time of the reform were not readily available. But according to the Federal Ministry of Family Affairs, Senior Citizens, Women and Youth, the share of female teachers in 2003/2004 in primary schools was 85.8 percent. In the lower and upper vocational and the academic track it was only 56.4, 61.8, and 50.2 percent, respectively. See <http://www.bmfsfj.de/doku/Publikationen/genderreport/1-Bildung-ausbildung-und-weiterbildung/1-4-Schulische-bildung/1-4-4-lehrkraefte.html>.

limited power to persuade parents into allowing their children to pursue an academic career that they perceive as incommensurate with their own social background, one would expect this particular policy change to *increase* the salience of parental background on educational outcomes.

To summarize, the reform intended to balance the advantages associated with creating homogeneous groups with an effort to dilute the influence of parental background. While it is somewhat unclear whether this was thought to be an important component of the reform, the tracking recommendation would now be made by teachers that were arguably in a better position to judge the suitability of students for subsequent tracks.

3 Conceptual framework

3.1 Peer effects and class homogeneity

Tracking may affect the efficiency of educational production as well as the variation in outcomes through altering the relationship between individual and peer quality.¹³ 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 of own education in parental 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.

While both peer effects and class homogeneity may affect the distribution of educational outcomes across parental background, it is important to keep in mind that tracks also differed in terms of resources devoted to it (see previous section).

3.2 Misallocation 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 misallocation of students to tracks. In particular,

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

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 standardized test scores between different school tracks (Baumert et al., 2003). Lehmann and Peek (1997) find that students with 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), and Dustmann et al. (2016) all find that students' exact birthday relative to an arbitrary enrollment cut-off date predicts track choice in Germany.¹⁴ While such findings are indicative of inefficiencies associated with tracking, both Jürges and Schneider (2011) and Dustmann et al. (2016) 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. (2016) 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.¹⁵ 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.

¹⁴Schneeweis and Zweimüller (2014) present evidence for Austria in which tracking also occurs at the age of ten.

¹⁵Jürges and Schneider relate this finding to "differences in verbal and non-cognitive skills at age 10."

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 (SOEP v30), a nationally representative survey carried out on an annual basis between 1984 and 2013. See [Wagner et al. \(2007\)](#) for a general description.

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, and includes all stages from primary to tertiary education.¹⁶ 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 birth¹⁷ and gender. We also always include a complete set of indicators for the size of the respondent’s locality during childhood¹⁸ and migrant status.¹⁹

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 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.²⁰ 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 occupation²¹ and the likelihood that respondents engaged in extracur-

¹⁶The 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.

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

¹⁸Categories are ‘no information available’, ‘city’, ‘large town’, ‘small town’, and ‘rural area’.

¹⁹Categories are ‘no information available’, ‘no migratory background’, ‘direct migratory background’, ‘indirect migratory background’, and ‘migratory background (not further specified)’.

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

²¹Conditional 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

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	12.25	12.06	12.45	12.25	12.16	12.36
Eligible for university	0.35	0.30	0.42	0.34	0.30	0.39
University degree	0.21	0.17	0.25	0.20	0.16	0.24
<i>Panel B. Individual background characteristics</i>						
<i>Panel B1. Basic characteristics</i>						
Year of birth	1962	1962	1961	1961	1962	1961
Educated parent	0.33	0.33	0.33	0.35	0.36	0.33
<i>Panel B2. Childhood place of residence was mostly...</i>						
...city.	0.20	0.20	0.21	0.12	0.13	0.12
...large town.	0.18	0.18	0.18	0.16	0.15	0.16
...small town.	0.22	0.22	0.21	0.22	0.22	0.21
...rural area.	0.36	0.35	0.36	0.47	0.47	0.48
<i>Panel B3. Migratory background</i>						
Direct	0.04	0.04	0.04	0.03	0.02	0.03
Indirect	0.12	0.12	0.11	0.07	0.07	0.07
Unspecified	0.00	0.00	0.00	0.00	0.00	0.00
No migratory background	0.84	0.84	0.85	0.90	0.90	0.90
<i>Panel B4. Sample information</i>						
Last observed in 2012	0.06	0.06	0.06	0.06	0.06	0.06
Observations	1,593	2,430	1,977	3,484	981	488

Statecodes as in notes to figure 1.

ricular activities at the age of 15 (such as having actively played a musical instrument or having done sports).²² 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 SOEP 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

respondents were 15 years old by a factor greater than four (three). These estimates are highly significant.

²²Conditional 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.

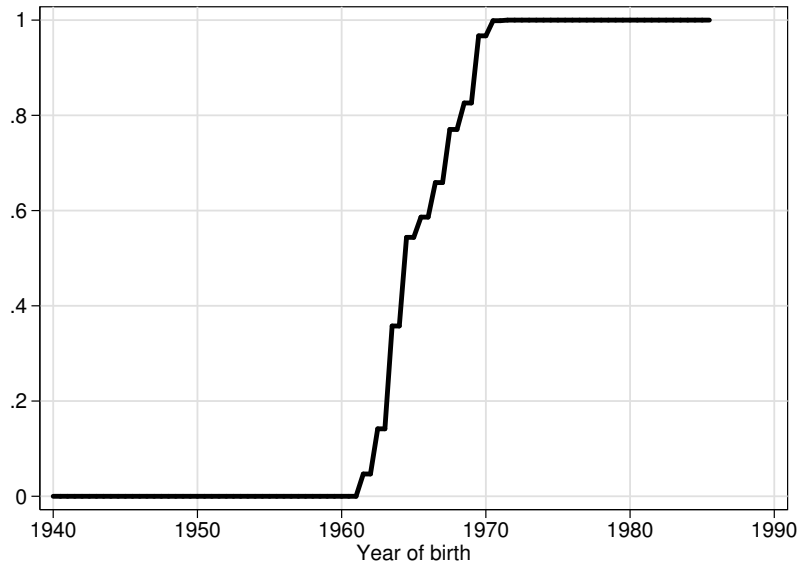


Figure 2: OS share against birth year in Lower Saxony.

these data allows us to calculate the percentage of students of one cohort that attended an OS school in a given school year.²³ 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 SOEP 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 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.

²³We ignore private schools which play a negligible role in Germany.

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.²⁴ In consequence, all individuals in our sample were born before 1985. We also exclude individuals born before 1950.

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

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

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).²⁵ α 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 of own education in parental education across states prior to the reform. We therefore always include state-background-fixed effects λ_s^b , where $b \in \{low, high\}$ indicates parental education. τ_c denotes cohort-fixed effects. All variables in \mathbf{x}_{isc}

²⁴Schooling 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.

²⁵Estimating 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.

are also interacted with an indicator of having educated parents. When the sample is pooled across males and females, we also interact λ_s^b , τ_c , and all variables in \mathbf{x}_{isc} with an indicator of respondents' gender. Finally, ϵ_{isc} is the usual white noise-error term.

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

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

where ψ_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 of own education in parental 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

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

The parameters of interest here are β_1 and β_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 ϕ_{sc}):

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

Note that the main effect of the reform is no longer identified and thus excluded here. β_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 requires that the gradient remains constant over time except for the effect of the reform. We relax the latter assumption by including ψ_c^b to allow for common changes in the gradient across cohorts:

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

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 whether or not prior the reform (1) trends in educational outcomes and gradients differed and (2) raw gradients differed between states.

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:

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

and

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

where all variables are defined as before and c is a linear cohort term. Testing whether our

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 (2)	Probit (3)	OLS (4)	Probit (5)	OLS (6)	OLS (7)	Probit (8)	OLS (9)	Probit (10)
<i>Panel A: Testing for common trends in outcomes prior to the reform.</i>											
$LS_s \times c$	0.001 (0.013)	0.003 (0.002)	0.014 (0.011)	0.001 (0.002)	0.007 (0.010)	0.007 (0.017)	0.001 (0.004)	0.005 (0.011)	0.000 (0.004)	-0.001 (0.013)	
Observations	3,295	3,295	3,292	3,295	3,292	3,143	3,143	3,140	3,143	3,143	
R-squared	0.161	0.124		0.100		0.144	0.096		0.100		
<i>Panel B: Testing for common trends in gradients prior to the reform.</i>											
$LS_s \times I[b = high] \times c$	-0.000 (0.042)	0.005 (0.007)	0.004 (0.023)	0.001 (0.005)	0.001 (0.022)	0.054 (0.044)	0.010* (0.006)	0.028 (0.017)	0.014* (0.008)	0.050* (0.027)	
Observations	3,295	3,295	3,232	3,295	3,036	3,143	3,143	3,128	3,143	3,114	
R-squared	0.195	0.155		0.140		0.181	0.131		0.140		

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).

estimates of ξ 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 ξ_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 ξ_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.

Results from estimating equations (4.6) and (4.7) are reported in panels A and B of table 2. While using a linear model to estimate the treatment effect in a DD specification is innocuous when the explanatory variables are all binary, including linear cohort-trends will result in these models being miss-specified ultimately as they will result in predictions outside the admissible unit interval. We therefore also report results from estimating probit-versions of (4.6) and (4.7). Standard errors clustered at the state-cohort-level are reported in parentheses in all tables that follow.²⁶

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 in gradient between Lower Saxony and other states may have widened prior to the reform: however, the point estimate suggests that the difference in gradients increased over ten years by only about half a year. The coefficients in columns (7) and (10) are also positive and even mildly significant. 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 education²⁷ 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 (columns (3) and (4a) and (4b)). The coefficient on the main effect for males (column (3)) is highly significant and suggests that males

²⁶We 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.

²⁷Results for university eligibility and university graduation are similar.

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	2.08*** (0.21)	1.08*** (0.41)	1.73*** (0.32)	2.56*** (0.33)	1.85*** (0.26)	1.79*** (0.35)
Educated parents × BV	-0.43 (0.32)	0.53 (0.51)	-0.83** (0.41)	-0.83** (0.40)	-0.93** (0.41)	-0.25 (0.48)
Educated parents × BW	0.08 (0.36)	0.73 (0.57)	-0.43 (0.45)	-0.36 (0.39)	-0.88** (0.40)	0.24 (0.47)
Educated parents × NRW	0.02 (0.30)	0.20 (0.50)	-0.10 (0.40)	-0.57 (0.38)	-0.73* (0.38)	0.03 (0.43)
Educated parents × RP	-0.44 (0.44)	-0.17 (0.56)	-0.10 (0.54)	-0.94* (0.53)	-0.97* (0.56)	-0.24 (0.54)
Educated parents × SH	-0.46 (0.42)	-0.49 (0.70)	-0.12 (0.66)	-1.01** (0.48)	-0.48 (0.56)	-0.83 (0.54)
<i>Joint significance of interactions.</i>						
<i>p</i> -value	0.51	0.36	0.33	0.18	0.16	0.43
Observations	3,295	3,295	3,295	3,143	3,143	3,143
R-squared	0.13	0.15	0.15	0.11	0.14	0.14

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.

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 0.18. Yet the estimate of the pre-reform gradient is significantly lower in Bavaria, Rhineland-Palatinate, and Schleswig-Holstein and the magnitude 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 (4a) and (4b), respectively. This shows that the overall differences are mainly driven by differences in the effect of maternal education on own education. Turning to columns (1) as well as (2a) and (2b), we find that differences are similar qualitatively yet less pronounced for females and mostly insignificant. Fathers' education may however have played a larger role in Bavaria.

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 further policy responses over and above the reform effort we study. In line with

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

	Pooled		Females		Males	
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.012 (0.121)	-0.009 (0.122)	0.050 (0.158)	0.028 (0.162)	-0.031 (0.196)	-0.050 (0.194)
Observations	13,577	13,577	7,023	7,023	6,554	6,554
R-squared	0.174	0.182	0.187	0.193	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 that are interacted with parental background. 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).

the notion that mothers are the primary care-givers in the absence of childcare, we show in a supplementary analysis in appendix C that persistent differences in pre-primary enrollment seem to account for 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 and, at the same time, differences in pre-reform gradients.

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 point 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*.

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 5. Results for the pooled sample are reported in panel A and results for females and males are reported in panels B and C, respectively.

Estimates 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 about one fifth of a year for individuals with uneducated parents. However, the

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

	Years of education			University eligibility (4)	University graduation (5)	Graduation conditional on eligibility (6)
	(1)	(2)	(3)			
<i>Panel A. Pooled sample.</i>						
Ref. × ed. parents	-0.520** (0.251)	-0.577** (0.275)	-0.846*** (0.291)	-0.092 (0.057)	-0.150*** (0.051)	-0.198* (0.104)
Reform	0.205 (0.158)					
Observations	13577	13577	13577	13577	13577	4790
R-squared	0.175	0.200	0.207	0.170	0.151	0.187
<i>Panel B. Females.</i>						
Ref. × ed. parents	-0.161 (0.309)	-0.135 (0.324)	-0.443 (0.354)	-0.006 (0.066)	-0.106 (0.066)	-0.148 (0.173)
Reform	0.112 (0.193)					
Observations	7023	7023	7023	7023	7023	2090
R-squared	0.187	0.212	0.219	0.177	0.147	0.199
<i>Panel C. Males.</i>						
Ref. × ed. parents	-0.949* (0.494)	-1.080** (0.495)	-1.301** (0.517)	-0.189** (0.084)	-0.201** (0.089)	-0.234* (0.134)
Reform	0.305 (0.243)					
Observations	6554	6554	6554	6554	6554	2700
R-squared	0.156	0.181	0.190	0.142	0.139	0.176
<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 that are interacted with parental background. 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).

estimate is not significantly different from zero. The effect on individuals with educated parents was negative and only marginally insignificant at the ten-percent level: the point estimate is $\widehat{\beta}_1 + \widehat{\beta}_2 = -0.309$. The estimate on the interaction term, -0.506 , is significant at the five-percent level suggesting that the gradient between parental and own education decreased by about half a year. Estimates of the effect on the gradient remain negative and significant after controlling for state-cohort-fixed effects (column (2)) and common shifts in the gradient over time (column (3)). Estimates in columns (4) and (5) show that the gradients in terms of the probabilities of completing the academic track and obtaining a university degree declined—the latter, which is also statistically significant, by more than ten percentage points.

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 gradient in years of education decreased by between 0.9 and 1.3 years and that all estimates in this case are significantly different from zero at the five-percent level. We will discuss plausible explanations for this gendered pattern in our results in section 7 below.

In column (6) we report estimates of the effect on the gradient in university completion *conditional on being eligible*. We obtain these estimates by restricting the sample to include only individuals who completed the academic track. The motivation behind this is to separate out the effect on university completion that comes about through a change in the subset of individuals who become eligible and the potential effect on the probability of transitioning from eligibility to university graduation. Note that the latter is more ambiguous: marginal students with educated (uneducated) parents that complete the academic track will likely have greater (lower) academic potential. Hence, one may expect the conditional gradient to widen after the reform. However, this is not what we observe in the data: for males only, the effect on the conditional gradient is estimated to be about -22 percentage points. This compares to an unconditional gradient of about 33 percentage points prior to the reform on average and more than 40 percentage points in Lower Saxony. In any case, a decomposition exercise in appendix B shows that about two thirds of the effect on the gradient in years of education for males is a result of the effect on the composition of individuals that complete the academic track, not on the conditional transition probability from eligibility to university graduation.

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 the data at the level of state-cohorts, the level at which the treatment variables varies. We then investigate whether our results are robust to alternative sample restrictions that allow us to rule out other potential explanations for our findings. In particular, we will rule out variation in pre-school enrollment, access to different types of schools, and the composition of families as potential alternative

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

	Females (1)	Males (2)
Reform	-0.489 (0.421)	-1.305*** (0.498)
R-squared	0.213	0.255
Observations	276	274

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.

explanations for our finding. 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 standard difference-in-differences equation. Second, it now becomes clear 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 1.3 years. The coefficient estimate for females carries the same sign but is much smaller in absolute terms and insignificantly different from zero. Reassuringly, standard errors on estimates are very similar to those reported in [table](#)

5.

6.2 Alternative samples

Recall that only one state, Lower Saxony, introduced the reform we focus on in this paper. It is therefore of great importance to rule out confounding factors. We do so in this section by carefully altering underlying samples. We rule out that our results are a consequence of differences in trends in pre-school enrollment, access to schools, and family composition. Table 7 reports results from estimating (4.5), the specification that includes state-background-, cohort-background-, and state-cohort-fixed effects, with our usual outcomes. Baseline estimates from column (3) of table 5 are reproduced in panel A for reference. We focus on the subsample of males throughout.

Differences in pre-school enrollment A potential confounding factor in our setting is pre-primary education. Pre-school enrollment has repeatedly been shown to affect both cognitive and behavioral outcomes, particularly among disadvantaged children.²⁸ 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 slots 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 graph 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.

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 capacity depends on its initial level, that is, if there are non-linear effects. This would be the case if, say, pre-school education is generally beneficial and the propensity to enroll varies with parental background ([Schütz et al., 2008](#)). We therefore ran a supplementary analysis based on matching the data plotted in figure 3 with our microdata (see appendix C). We show that our data are consistent with a u-pattern.²⁹

²⁸[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.

²⁹This pattern is consistent with a higher propensity for uneducated parents to have their children attend

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

	Years of education (1)	Eligibility (2)	University degree (3)	Degree conditional on eligibility (4)
<i>Panel A. Baseline.</i>				
Baseline	-1.301** (0.517)	-0.189** (0.084)	-0.201** (0.089)	-0.234* (0.134)
Observations	6,554	6,554	6,554	2,700
<i>Panel B. States with and without similar pre-primary capacity.</i>				
LS, NRW, SL	-1.264** (0.521)	-0.190** (0.086)	-0.195** (0.089)	-0.189 (0.134)
Observations	3,364	3,364	3,364	1,389
LS, BW, RP, and BV	-1.274** (0.559)	-0.189** (0.091)	-0.200** (0.095)	-0.275* (0.146)
Observations	4,147	4,147	4,147	1,685
<i>Panel C. Rural vs. urban upbringing.</i>				
Rural childhood	-1.261** (0.638)	-0.190* (0.103)	-0.223** (0.103)	-0.199* (0.105)
Observations	3,845	3,845	3,845	3,056
Urban childhood	-1.969** (0.833)	-0.208* (0.106)	-0.264* (0.145)	-0.378** (0.171)
Observations	2,432	2,432	2,432	1,732
<i>Panel D. Family background.</i>				
Mother's age at birth ≥ 20	-1.263** (0.544)	-0.191** (0.089)	-0.190** (0.095)	-0.203 (0.139)
Observations	6,206	6,206	6,206	2,619
Mother and father present	-1.415** (0.584)	-0.185** (0.093)	-0.240** (0.095)	-0.257** (0.129)
Observations	5,637	5,637	5,637	2,389
No migratory background	-1.466*** (0.518)	-0.229*** (0.087)	-0.215** (0.092)	-0.188 (0.139)
Observations	5,597	5,597	5,597	2,340
<i>Panel E. Excluding partially treated cohorts.</i>				
$OS_c = 0$ or $OS_c = 1$	-1.387** (0.611)	-0.173* (0.101)	-0.255** (0.099)	-0.383*** (0.139)
Observations	4,636	4,636	4,636	1,825
<i>Panel F. Excluding cohorts born in the second half of 1968 and later.</i>				
Born before summer of 1968	-1.943*** (0.729)	-0.257* (0.134)	-0.302** (0.141)	-0.252 (0.209)
Observations	4,507	4,507	4,507	1,791

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 state-cohort-, state-background-, and cohort-background-fixed effects as well as dummy variables for the size of respondents' childhood place of residence and migrant status that are interacted with parental background.

pre-school educational institutions. It is the opposite of the pattern that has been found recently by [Schütz et al.](#) who base their analysis on recent data from the *Third International Mathematics and Science Study* (TIMMS) and its replication for a partly different set of countries (TIMMS-Repeat). 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.

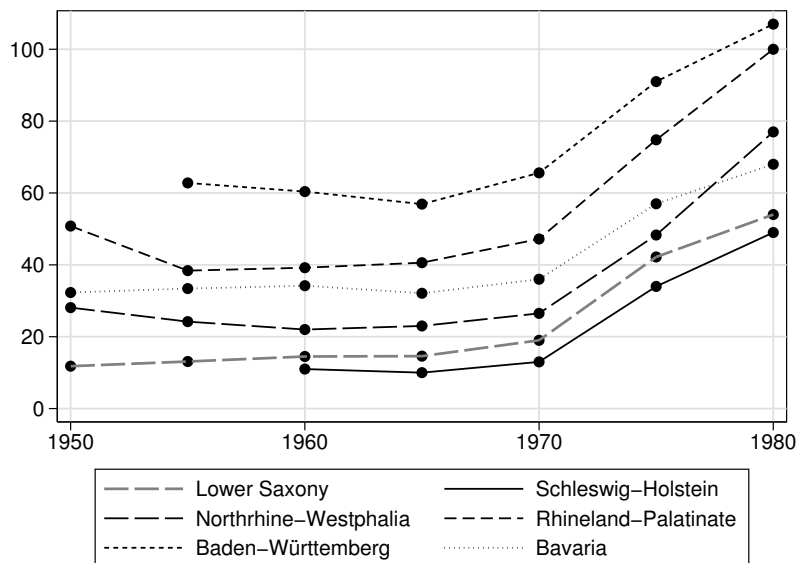


Figure 3: Pre-school capacity (slots per 100 children between three and five). *Source:* Erning et al. (1987, p. 37).

Intriguingly, we also find that differences in pre-school capacity across states explain differences in the gradient we encountered in section 4.3.³⁰

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 panel B of table 7, we split the sample of control states according to whether levels and trends in pre-primary capacity have been similar to those in Lower Saxony. We exclude Bavaria (BV), Baden-Württemberg (BW), and Rhineland-Palatinate (RP) for a sample that compares Lower Saxony to states that have had very similar pre-primary capacity, namely Northrhine-Westphalia (NRW) and Schleswig-Holstein (SH) (see figure 3). We also test using excluded states as the exclusive set of control states. We find that it does not matter whether or not states that were similar in terms of trends and levels in pre-primary enrollment constitute the control group.

Changes in access 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, there may still be differences across states. Lower Saxony, in particular, may be different as it is thinly populated in comparison to other states.³¹ For instance, Kramer (2002) documents an increase in access to upper vocational and academic track schools especially in

³⁰Since 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.

³¹It is the second-largest state of Germany in terms of area yet only the fourth-largest in terms of inhabitants.

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 and small towns gaining access to upper vocational and academic track schools, we would find no effect of the reform once we exclude these students. We therefore split the sample according to whether individuals grew up in large cities and towns *vis-à-vis* small towns and rural areas (panel C). We find that the point estimate does not change when we restrict the sample to individuals that grew up in rural areas. We find a much larger effect on the gradient, however, when we investigate individuals that grew up in urban areas. However, the standard error also increases in this case. We conclude that there is very clear evidence for a decrease in the gradient in both subsamples.

Family background Next, we check whether differences in trends in early pregnancy, ‘broken families’, and migrant population may act as a confounder. [Bertrand and Pan \(2013\)](#), for instance, find that single mothers invest fewer resources into their sons. If trends in the proportion of individuals with single mothers differed systematically between states and if having a single mother is correlated with parental education, our estimates may be reflecting these changes rather than the effect of the reform. We restrict our sample to individuals whose mothers were at least 20 years of age when they were born, individuals with both biological parents present, and individuals without migratory background. Our findings indicate that our results are robust, i.e. that trends in these variables have not differed between states.

Excluding partially treated cohorts The introduction of OS schools in Lower Saxony took almost one decade and during this time period selection may have played a role. One indication of this would be if results changed once cohorts that were only ‘partially treated’ in the sense that the roll-out of OS schools had started but was not completed at the time they were in the respective age. As is clear, this affects cohorts born during the 1960s. 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 that sense, this test is not a robustness check in which we aim to eliminate an alternative explanation for our results. In any case, the co-existence of two tracking regimes potentially produces outcomes that differ from those under complete de-tracking.³² 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.

³²Recall that this is why Hesse was excluded, a state in which the two systems exist next to each other.

Binding recommendations A final test concerns the effects of the abolishment of binding tracking recommendations that we discussed in section 2. Tracking recommendations in Lower Saxony were binding before the school year 1979/80 and non-binding afterwards. However, under both regimes did parents have the option to send their children to a lower track. A number of studies for Germany have documented that if parents have the choice, they tend to overrule recommendations in concordance with their own educational attainment, i.e. educated parents tend to reject track recommendations in favor of sending their children to the academic track while the opposite is true for uneducated parents (e.g. [Ditton et al., 2005](#); [Harazd and van Ophuysen, 2008](#)). [Lohmann and Groh-Samberg \(2010\)](#) find that rejecting track recommendations in favor of the academic track pays off for students with educated parents in that they are more likely to attend the academic track at the age of 17. Assuming that the reform was effective in reducing the effect of parental education on own education, an assumption supported by results discussed previously, it is therefore reasonable to assume that the abolishment of binding recommendations in Lower Saxony by 1980 may have reversed some of the positive effects the reform may have had on equality of opportunity.

This is indeed what we may tentatively conclude from results in panel F of table 7 in which we drop all individuals that were born in the second half of 1968 or later and that were thus affected by non-binding recommendations. The point estimate on the interaction term of interest indicates that the gap may have narrowed by as much as 1.9 years for males. This would suggest that the negative effect of de-tracking on the gap was particularly pronounced during the first years of the reform. However, while the estimate is significant at the one percent-level, it also carries a standard error that is about 50 percent greater than in our baseline estimation. These results are also in line with previous results on the effect of a change to a regime of non-binding recommendations in Lower Saxony ([Roth and Siegert, 2015](#)).³³

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

³³[Roth and Siegert \(2015\)](#) document an attenuation of the effect of parental education on the probability of attending the academic track at the age of 13 or 14 in Lower Saxony already for cohorts for which the recommendation was still binding.

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	-0.473 (0.565)	-0.464 (1.016)	0.462 (0.670)	0.417 (1.197)
Observations	21,423	21,423	19,999	19,999
R-squared	0.13	0.14	0.13	0.14

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 that are interacted with parental background.

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 ten³⁴, 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 introduction of de-tracking, it would seem that either unobserved factors are driving our results or that we are picking up random quirks in the underlying sample. The present placebo test allows us to rule out these possibilities to some extent.

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 clustered around the individual-identifier for POLS estimates. Results are reported in table 8 for females and males.³⁵

³⁴While 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).

³⁵As 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).

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.

7 Discussion

The results presented in the previous section 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 one and one and a half 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 $(0.305 - 0.949) \times 0.07 \approx -4.5$ percent for males with educated parents and $0.305 \times 0.07 \approx 2.1$ percent for males with uneducated parents. While these effects would be very large, particularly for a reform that affected only two mandatory grades within the education system, they are smaller than those found by [Meghir and Palme \(2005\)](#) who report estimates of -7.7 percent and 3.1 percent for males, respectively.^{36,37} 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, it may still be a lower bound for the true effect: first, our matching of individuals to states in which they received schooling will be erroneous to some extent and this mismeasurement 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

³⁶However, we found that the effects we calculate here would be too small to be picked up in an empirical analysis (along the lines of the analysis presented in section 6.3) given our sample.

³⁷Remember, however, that the Swedish reform was very different in that tracking occurred only after six years prior to the reform. It also encompassed additional elements such as transfer payments and an increase in compulsory schooling that may have affected wages.

(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 an informed choice on behalf of their children, keeping these cohorts in the estimation sample will likely tend to bias our estimate towards zero. However, the estimated effect on the gradient is only marginally greater when we exclude partially treated cohorts from the estimation sample (see section 6).

The gendered pattern we find in the data suggests four non-exclusive hypothesis that, unfortunately, cannot be addressed given the nature of the experiment.³⁸

1. In-school peer effects are more important for boys than for girls. If parental education were positively correlated with attributes that increase academic potential in students, the reform on average would have resulted in a decrease in peer quality for high-quality students and *vice versa*.
2. Teachers effects are more important for boys than for girls. Recall that academic track teachers were of higher quality—at least when judged by pay and training (see section 2.1)—and teachers in orientation stage schools came from all three traditional secondary tracks.
3. Primary school teachers in Germany were predominantly female while the gender ratio of teachers in secondary schools was often more balanced (section 2.2). Hence, a third hypothesis is that criteria for recommendations for males change as a result of an increase in the share of male teachers.
4. There are gendered patterns in skill formation that interact with the timing of tracking. More precisely, the *observability* of certain skills relevant for tracking that have less of a relationship to parental background become comparatively more salient only at a later point for boys compared to girls. For instance, there is convincing evidence of gendered differences in brain development (e.g. [Lenroot et al., 2007](#)). And while gendered differences in non-cognitive skills and reading tend to be fairly stable across age ([Bertrand and Pan, 2013](#)), male advantage in math tends to increase with age ([Fryer and Levitt, 2010](#); [Bertrand and Pan, 2013](#)).

Clearly, alternative experiments, natural or controlled, are needed to distinguish between the above hypotheses. These require random variation in peer composition, teacher allocation in terms of quality and gender, and the timing of tracking, respectively. As the reform we study here affected all of the above, peer quality, teacher quality and gender, and the age of tracking, and, due to the retrospective nature of our data, we lack information about precise circumstances at the time of tracking, there is no way of disentangling these channels.

³⁸Note that this finding may not have been surprising had we found that there was a much smaller gradient for girls to begin with. However, the gradient prior to the reform was more than two years, approximately 80 percent of the gradient for males (see section 4.3.2).

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 as measured by the effect of parental education on own education. 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 alternative specifications and samples. There is no indication that the reform had any effect on females—a finding that suggests interesting alleys 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.
- Bertrand, M. and Pan, J. (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics*, 5(1):32–64.
- 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.
- Buis, M. L. (2015). Not All Transitions Are Equal: The Relationship Between Effects on Passing Steps in a Sequential Process and Effects on the Final Outcome. *Sociological Methods & Research*.
- 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. 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 Bohnkamp, 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.
- Ditton, H., Krüskens, J., and Schauenberg, M. (2005). Educational Disparities and the Contribution of Family and School. *Zeitschrift für Erziehungswissenschaft*, 8(2):285–304.
- 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. (2016). The Long-Term Effects of Early Track Choice. *Economic Journal*.

- 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.
- Fryer, R. G. J. and Levitt, S. D. (2010). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*, 2(2):210–240.
- 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 Pilot 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.
- Harazd, B. and van Ophuysen, S. (2008). Reasons for Choosing a Non-Recommended School Type. *Zeitschrift für Erziehungswissenschaft*, 11(4):626–647.
- 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 Schulentwicklung 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., Pasheran, 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.
- Lohmann, H. and Groh-Samberg, O. (2010). Acceptance of Secondary School Track Recommendations and Their Effects on Educational Achievement. *Zeitschrift für Soziologie*, 39(6):470–492.
- 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.
- 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.
- Roth, T. and Siegert, M. (2015). Freedom versus Equality? The Effect of Compusory Primary School Recommendations on Social Inequality in Secondary Education. *Zeitschrift für Soziologie*, 44(2):118–136.
- 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: 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.	0.03	0.03	0.03	0.03	0.03	0.03
...state of last school visit.	0.69	0.69	0.69	0.71	0.72	0.70
...childhood state.	0.15	0.15	0.16	0.14	0.14	0.15
...first state in which observed.	0.13	0.13	0.12	0.12	0.11	0.13
Observations	11,567	5,970	5,597	2,010	1,053	957

Statecodes as in notes to figure 1. Pre-primary capacity is the number of slots available per 100 children between three and five. Raw data from [Erning et al. \(1987\)](#) and partly interpolated (see text).

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 Decomposition of the effect on years of education by degrees

The effect of the reform on the gradient in university graduation we report in section 5 could be driven by both effects on the composition of eligible individuals as well as effects on the probability of graduating from a university *conditional on being eligible*. Also, it is of interest to gauge to which extent the effect on the gradient in years of education, our main variable of interest, is driven by changes in completion probabilities. Our data allow us to decompose the effect of the reform on years of education for individuals with and without educated parents into the contributions from the reform's effect on the probabilities of making the transitions from non-eligibility to eligibility for university studies (that is, finishing the academic track) and from university eligibility to university graduation.³⁹

Denote by p_1 the probability of transitioning to university eligibility and by b_1 the effect of the reform on this probability. Similarly, denote by p_2 the probability of transitioning from eligibility to a university degree and let b_2 be the effect of the reform on this conditional probability. Let l_0 , l_1 , and l_2 be average years of education associated with the three exclusive outcomes non-eligibility, eligibility without having earned a degree, and having earned a degree. The approximate effect of the reform on expected years of education $E[Y]$ can then be calculated as

$$\frac{\partial E[y]}{\partial D} = \{1 \times [(1 - p_2)l_1 + p_2l_2 - l_0]\}b_1 + \{p_1 \times [l_2 - l_1]\}b_2. \quad (\text{B.1})$$

See, for instance, [Buis \(2015\)](#). Another way of writing the above equation is

$$\frac{\partial E[y]}{\partial D} = \left\{ \begin{array}{ll} 1 & \times p_1(1 - p_1) \\ & \times [(1 - p_2)l_1 + p_2l_2 - l_0] \end{array} \right\} \lambda_1 + \left\{ \begin{array}{ll} p_1 & \times p_2(1 - p_2) \\ & \times [l_2 - l_1] \end{array} \right\} \lambda_2, \quad (\text{B.2})$$

where λ_1 and λ_2 denote log odds ratios. From (B.2) we see that the effect on years of education can be decomposed in changes in transition probabilities and that these enter with weights which are denoted in curly brackets. These weights depend on (1) the proportion at risk (in the case of the first transition this is unity), (2) the differentiating capacity of the k th transition as captured by $p_k(1 - p_k)$, and (3) the difference between expected outcomes.

³⁹Our data do not allow us to identify failed attempts to earn university degrees.

Estimates of l_0 , l_1 , and l_2 are simply obtained from the data by averaging within each group. We restrict the sample to individuals that received schooling in Lower Saxony and were unaffected by the reform. Estimates are $\widehat{l}_0 = 10.59$, $\widehat{l}_1 = 12.61$, and $\widehat{l}_2 = 16.27$. One minor caveat of this analysis is that we need to assume these averages are fixed whereas, in fact, they may change to some extent as a consequence of the reform.⁴⁰

Estimates of b_1 and b_2 are obtained from standard DD-type regressions for males with educated and uneducated parents separately, e.g.

$$P(\textit{eligibility}_{isc} = 1) = b_1 D_{sc} + \mathbf{x}'_{isc} \gamma^1 + \sigma_s^1 + \tau_c^1 + \epsilon_{isc}^1, \quad (\text{B.3})$$

and

$$P(\textit{graduation}_{isc} = 1 | \textit{eligibility}_{isc} = 1) = b_2 D_{sc} + \mathbf{x}'_{isc} \gamma^2 + \sigma_s^2 + \tau_c^2 + \epsilon_{isc}^2, \quad (\text{B.4})$$

where σ^1 and σ^2 and τ^1 and τ^2 denote state- and cohort-fixed effects, \mathbf{x}_{isc} includes the same controls as before, and ϵ^1 and ϵ^2 are white noise error terms. In order to obtain an estimate of the transition probability p_2 , the probability of obtaining a degree *conditional on being eligible*, we restrict the sample to eligible individuals when estimating (B.4).

Panel A of table 10 reports baseline estimates of the effect of the reform on years of education (columns (1) and (4)). Reassuringly, the latter suggest an effect on the gradient of $-0.795 - 0.351 = -1.146$, very close to the estimates we obtained in section 5. It also reports estimates from simple LPMs⁴¹ of b_1 and b_2 (columns (2) and (5), and (3) and (6), respectively). These estimates are in line with our expectations in that they carry negative signs for transitions of individuals with educated parents and positive signs for transitions of individuals with uneducated parents. Finally, estimates of the baseline probabilities are obtained simply by taking means from the data and reported in panel B of the same table along with our decomposition results. Again, we use only observations on males from Lower Saxony that were not affected by the reform.

The decomposition (B.1) is on approximative but the approximation seems very close. Comparing the total effects between panels A and B we observe small differences between the estimated coefficients we obtain from (B.1) and those we obtain from the actual econometric model given by (B.3) and (B.4). These are likely due to changes in the outcomes and transition probabilities over time. We further observe in panel B that the transition probabilities are much higher for individuals with educated parents: prior to the reform the odds of completing the academic track for individuals with educated parents that went to school in Lower Saxony at the age of ten were almost 2:1. Conditional on having completed the academic track their odds were better than 4:1 to also obtain a university degree. This compares to odds of about 0.3:1 and 1:1 for

⁴⁰We find minor differences between averages of our estimates of l_0 and l_2 by parental education that are statistically significant. This suggests, for instance, that conditional on having obtained a university degree, individuals with educated parents are more likely to have obtained degrees that take more time to complete. However, we found that effects obtained from (B.1) matched estimated effects more closely when we used estimates for the outcomes obtained after pooling individuals with educated and uneducated parents.

⁴¹We found that marginal effects from a sequential logit model are virtually identical.

Table 10: Effect of reform on probability of transition from non-eligibility towards eligibility for university studies and from eligibility towards a university graduation by parental background.

	Educated parents			Uneducated parents		
	Years of education (1)	Eligibility (2)	Graduation conditional on eligibility (3)	Years of education (4)	Eligibility (5)	Graduation conditional on eligibility (6)
Reform	-0.795* (0.410)	-0.114* (0.063)	-0.067 (0.081)	0.351 (0.242)	0.040 (0.041)	0.175* (0.097)
Observations	2141	2141	1327	4413	4413	1373
R-squared	0.048	0.034	0.058	0.035	0.028	0.078
<i>Panel B. Decomposition.</i>						
	Calculated effect ^a			Calculated effect ^a		
Estim. transition probabilities ^b		0.64	0.84		0.248	0.500
Abs. contribution	-0.724	-0.560	-0.165	0.317	0.154	0.163
Rel. contribution	1	0.773	0.227	1	0.485	0.515

Standard errors clustered at the state-cohort-level in parentheses. *, **, and *** denote significance at the ten-, five-, and one-percent level, respectively.

^a Based on (B.1).

^b Averages of subsample of cohorts in Lower Saxony that were not affected by the reform.

individuals with uneducated parents.

Our results suggest that the negative effect of the reform on years of education for males with educated parents was a consequence mostly of a decrease by 11.4 percentage points in the probability of graduating from the academic track. The effect of the reform on this first transition accounts for more than three fourths of the decrease in years of education for this group. The positive effect on years of education for males with uneducated parents is not significantly different from zero (p -value of 0.147) and is small in comparison. Only the effect on obtaining a university degree conditional on being eligible is sizable at 17.5 percentage points and significantly different from zero at the ten-percent level. The effect on becoming eligible is small by comparison at four percent. Nevertheless, the contribution of the latter to the effect on years of education is still almost half for this group, mainly because all individuals are initially at risk of completing the academic track.

Taken together, results in table 10 indicate that approximately $(-0.560 - 0.154)/(-0.795 - 0.317) = 64$ percent of one year of the decrease of the change in the gradient is attributable to changes in the probability of becoming eligible for university. In other words, we would still observe a decrease in the gradient by about three to four fifths of a year if the reform had no effect on the transition probabilities p_2 . While plausible, this result remains somewhat speculative as standard errors on some of the parameters required for this exercise are fairly large.

C Differences in access to pre-primary across states 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 pre-school capacity between states around the 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 gradients and pre-primary capacity, the number of slots available per 100 children aged 3-5 (as a proxy for enrollment in pre-primary) and equality of opportunity in our data. We first show that differences in capacity account for differences in the gradient prior to the reform (as reported in section 4.3). More importantly, we find that in contrast to the findings in Schütz et al. (2008), the relationship between the gradient and capacity follows a u-shape in our data. While this would suggest that changes in capacity may be confounding our results after all, we show in section 6 that our main result is robust to restricting the set of control states to states those similar to Lower Saxony in terms of slots available.

We proceed by collapsing the data at the state-cohort-level (as in section 6.1) and regressing gradients on state- and cohort-fixed effects as well as capacity and capacity squared. Again, the number of individuals in each state-cohort is used as regression weight. We use the data depicted in figure 3 (obtained from tabulations in Erning et al. (1987)) and impute missing in-between

Table 11: Effect of pre-primary capacity (PPC) on the gradient between parental and own education for males.

	(1)	(2)	(3)	(4)
Constant	3.16*** (0.40)	5.04 (3.09)	4.24*** (0.87)	3.35*** (0.42)
State: BV	-1.53*** (0.52)	-0.18 (2.62)		
State: BW	-1.77*** (0.57)	-1.03 (4.66)		
State: NRW	-1.01** (0.48)	-0.13 (1.62)		
State: RP	-1.58** (0.66)	-0.11 (3.21)		
State: SH	-1.44 (1.13)	-1.81 (1.30)		
Pre-primary capacity / 100		-16.44 (26.44)	-11.42** (5.30)	-5.96*** (1.81)
(Pre-primary capacity / 100) ²		20.13 (33.66)	11.15 (6.96)	5.56*** (1.68)
<i>Results of F-tests for joint significance (p-values).</i>				
State-FEs	0.03	0.78		
Year-FEs				
Capacity variables		0.82	0.01	0.00
R-squared	0.15	0.15	0.12	0.07
Observations	80	80	80	153

Asymptotic standard errors in parentheses. Observations weighted by the number of individuals in each cohort-state. All regressions include sets of cohort- and state-fixed effects.

observations by linear interpolation. As pre-primary schools (that is, kindergartens in Germany) usually take in children between the ages of three and five, we associate each birth cohort with the available capacity in their respective states five years after their year of birth.

Results are reported in table 11. In the first column we regress gradients only on set of state-fixed effects with Lower Saxony as the excluded group and a constant that picks up the gradient in Lower Saxony. These variables will always be included. As we are interested in explaining differences in gradients across states prior to the reform, we exclude observations on cohorts born early enough to not be affected. We find a pattern similar to the one reported in table 3 in 4.3, that is, Lower Saxony's gradient is positive and significantly different from zero while those in other states are significantly smaller. These differences disappear mostly once we control for the capacity rate and the capacity rate squared (column (2)) and hence our claim that differences in pre-primary capacity may account for differences in the pre-reform gradient across states. The coefficients on capacity and capacity squared suggest an inverted-u-shaped relationship between

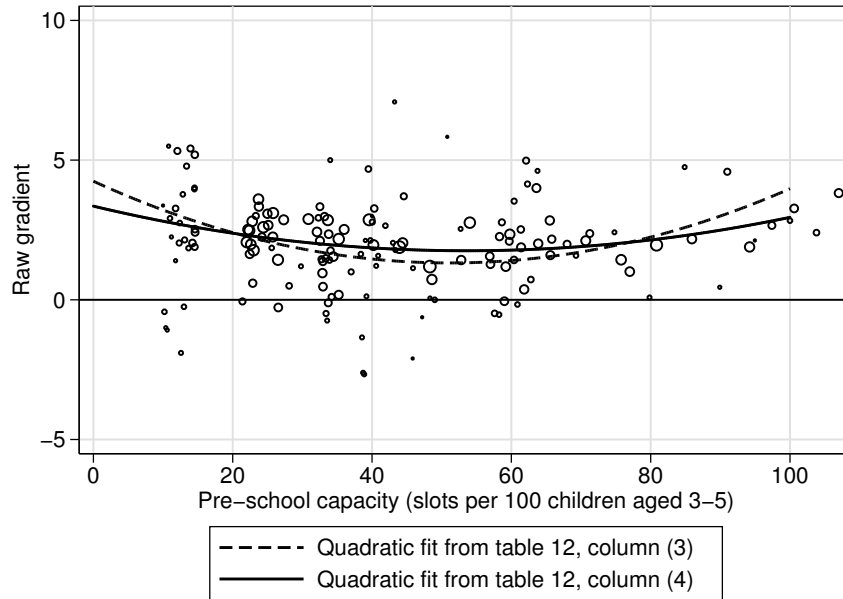


Figure 4: Illustration of the effect of pre-primary capacity rates on the gradient between parental and own education. Based on results in columns (3) and (4) of table 11. Only cohorts that went to school in Lower Saxony after the implementation of the reform are excluded. The size of the circles is proportional to the number of observations in the respective state-cohort.

the gradient and pre-primary education. However, both the capacity variables and the state-fixed effects are individually and jointly insignificant. When we exclude state-fixed effects in column (3) the relationship between the gradient and pre-primary education remains and the proxy variables and its square are jointly significant. The relationship is not altered when we run the regression on all observations with the exception of state-cohorts from Lower Saxony that were affected by the reform (column (4)). Figure 4 uses results from columns (3) and (4) to visualize this relationship.

Recall that [Schütz et al. \(2008\)](#) find an inverse-u-shaped pattern for their sample of younger cohorts in mostly developed countries while our results for a sample of individuals born mostly in the 1950s, 1960s, and 1970s indicate the opposite, a u-shaped pattern. An inverse-u would be consistent with the notion that pre-primary schooling is generally beneficial for children in terms of later educational attainment and that the propensity to enroll children is higher among educated parents and *vice versa*. Given that cohorts that had children born in the 1950s and 1960s in Germany experienced near-full employment while lower-class parents more recently are more likely to be unemployed or to forgo enrolling their children for other reasons. Hence, differences in our results may be attributed to changes in the relative propensity to enroll children.

The finding in this appendix 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 does not affect our estimates.