The Effect of Preschool Attendance on Secondary School Track Choice in Germany Evidence from Siblings

Martin Schlotter

Ifo Working Paper No. 106

July 2011

Ifo Working Paper No. 106

The Effect of Preschool Attendance on Secondary School Track Choice in Germany Evidence from Siblings*

Abstract

We study the effect of preschool attendance on secondary school track choice in Germany which is a crucial outcome that largely predicts educational pathways of children. Using data from the German Socio-Economic Panel, multivariate models show a significant positive association between years of preschool attendance and a child's probability of attending the highest school track, that is, German Gymnasium. Including family fixed effects in a sibling model, our estimates become considerably smaller and are no longer significant, indicating an upward bias in multivariate models. Accounting for several sibling-specific covariates, such as measures of innate ability and social skills, does not change this result. The low intensity of Germany's center-based preschool system might be a reason for the zero effects.

JEL Code: I2, J13, J24.
Keywords: Preschool education, sibling models, GSOEP.

Martin Schlotter
Ifo Institute for Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
Phone: +49(0)89/9224-1226
schlotter@ifo.de

We thank Silke Anger, Oliver Falck, Guido Schwerdt, C. Katharina Spieß, Ludger Wößmann, as well as seminar participants at Ifo Munich, IAW Tübingen, DIW Berlin, and participants of the Royal Economic Society 2011 Annual Conference for helpful discussions and comments. All remaining errors are our own.

1 Introduction

It is well-known that early childhood is a crucial period for the development of future socio-economic outcomes. Several studies show that living conditions in early life are valid predictors of later individual development in many dimensions, including health and education (see, for example, Maccini and Yang 2009; Gould et al. 2011). This is particularly true in the case of early differences in human capital endowment. In line with theoretical considerations by James Heckman, who models the skill-formation process over the lifecycle (see, for example, Cunha and Heckman 2007; Heckman et al. 2006), many empirical studies find that early investment in human capital has positive medium- and long-term effects on adolescents' and adults' socio-economic outcomes and yields higher returns than later remediation measures. This result is mainly based on several experimental studies such as the Perry Preschool Program (see Belfield et al. 2006; Heckman et al. 2010, for a follow-up) or the Abecedarian Project (see Temple and Reynolds 2007), as well as on quasi-experimental evidence from larger programs such as Head Start (see Garces et al. 2002; Ludwig and Miller 2007).

This paper analyzes the causal effect of years of preschool attendance in Germany on the probability of attending the highest secondary school track after primary school, that is, high school (Gymnasium). Secondary school track attendance is a widely studied outcome in many German analyses since it largely predicts future outcomes of children such as cognitive skills, secondary school degree, and earnings (Dustmann 2004; Jürges and Schneider 2007; Baumert et al. 2009). We use a sibling approach in which we include family fixed effects and therefore capture all unobservable effects at the family level that affect siblings' preschool attendance and secondary school track choice alike. In contrast to existing cross-sectional studies and our own cross-sectional estimates, we find that an additional year of preschool attendance has no significant effect on the probability of attending high school when we account for family-level heterogeneity.

This is a surprising finding—at least initially. However, characteristics of the German early childhood education system may explain it. First, the German preschool system is center based and very different from many experimental settings or other model programs such as Head Start. These other programs often include nutrition programs, substantial involvement of parents, and home visits. Second, studied U.S. programs are mostly targeted at children in need, whereas German preschool is a large-scale institution open to all children between the ages of three and six. In fact,

German preschool more closely resembles universal prekindergarten programs (Pre-K) in the United States or similar programs in other countries.[1]

Our sibling approach goes beyond finding partial correlations between preschool attendance and secondary school track choice, thereby complementing multivariate analyses that were carried out in other German studies on the topic (see Spiess et al. 2003; Büchner and Spiess 2007; Landvoigt et al. 2007; Seyda 2009). These studies show significant positive associations between overall preschool *attendance* and the probability of attending the highest secondary school track, that is, German high school (Gymnasium) (see Landvoigt et al. 2007), and, for immigrant children, a negative relationship between preschool attendance in the year prior to school entrance and the probability of attending the lowest secondary school track, that is, German basic school (Hauptschule) (see Spiess et al. 2003). Studies that identify significant effects of the *years* spent in preschool include Büchner and Spiess (2007), who find a negative association with the probability of attending basic school, and Seyda (2009), who finds a positive correlation with the probability of attending high school.

Since preschool attendance is not randomly assigned to children and might depend on many (non-observable) factors, causation is a central issue in interpreting preschool effects in non-experimental programs. In this paper, we use a sibling approach to tackle the causality problem. The German Socio-Economic Panel (GSOEP), which samples households and provides information on all household members, allows us to study whether variation in years of preschool attendance across siblings leads to differences in the probability of attending the highest secondary school track. Econometrically, this means estimating family fixed effects models in which we can capture all unob-

[1]Several papers study the effects of universal Pre-K and similar center-based programs on cognitive outcomes in the United States. Gormley et al. (2005) find positive short-term effects of universal Pre-K in Oklahoma on student test scores, in particular for Hispanics and blacks. Loeb et al. (2007) finds positive effects of Pre-K on test scores in U.S. kindergartens with special effects for Hispanics. Magnuson et al. (2007) show a positive correlation between Pre-K and test scores in U.S. kindergartens; however, the effects already fade after the first grade of primary school. Fitzpatrick (2008) uses a difference-in-differences approach to evaluate the Pre-K program in Georgia (USA) and shows positive effects on disadvantaged children's academic achievement in reading and math at grade 4. Similarly, Cascio (2009) and Dhuey (2011) estimate the long-term effects of introducing kindergarten into public schools in the 1960s and 1970s. Although U.S. kindergartens are part of the U.S. K-12 school system, at that time they were very similar to preschool institutions. As a consequence of the expansion of public kindergartens, these studies find very modest effects on reduced school drop-out rates and less institutionalization among whites (Cascio 2009), as well as negative effects on being below grade and positive impacts on wages for Hispanic children (Dhuey 2011). Outside the United States, Berlinski et al. (2009) find positive short-term effects of preschool expansion on test scores in primary schools in Argentina with bigger effects in more disadvantaged municipalities. Havnes and Mogstad (2011) find a positive long-term impact of child care expansion in Norway on individual outcomes such as educational attainment and labor market participation. Their effects are mostly driven by benefits for children of low educated mothers.

served family-level factors that have the same linear, additive effect on both preschool attendance and later outcomes of the siblings.

This method is frequently employed in the economic literature dealing with causation issues in econometric identification.[2] For example, and with relevance to our own work, Garces et al. (2002) use a sibling approach to identify long-term effects of Head Start participation on outcomes such as high school completion and college attendance; Berlinski et al. (2008), in a study of Uruguay, exploit within-family variation in preschool attendance across siblings to show positive effects of preschool education on years of completed education and negative effects on school drop-out.

In standard OLS estimations we find significant positive associations between one additional year of preschool attendance and the probability of attending high school, varying in magnitude between sub-samples of children with different socio-economic background. However, when we include family fixed effects in a sibling sample, the impact of an additional year of preschool attendance becomes significantly smaller and is not significantly different from zero. Nevertheless, we do find some evidence that children with migration background whose mothers do not speak German benefit more from an additional year of preschool than children with migration background having a German-speaking mother. Several covariates account for sibling-specific differences (affecting within-family variation in years of preschool attendance and track choice at the same time), such as birth order, age at school entrance, household income, and mothers' employment at the start of preschool. In addition, the GSOEP provides data from a test assessing innate ability of adolescents, as well as information on the social skills of three year old children (reported by mothers). These data allow us to investigate whether within-family differences across siblings in these dimensions affect preschool attendance and our outcome at the same time. We find that neither sibling-specific innate ability nor social skills is likely to explain within-family differences in preschool. This makes us confident that we have found causal effects of years of preschool attendance.

Several conclusions can be drawn from these findings. Years spent in preschool seems to be highly selective so that standard multivariate OLS models may not capture all crucial factors that affect a family's preschool decisions. Non-significant effects in specifications with family fixed effects show that non-targeted, low-intensity programs such as the German preschool might have no effect on educational outcomes in the medium-term. This is in line with the results from most other universal preschool programs that find, if any, very modest overall effects and/or benefits that are confined to disadvantaged children. For example, German preschool does not explicitly focus

[2]See Griliches (1979) for an early overview that deals with advantages and caveats of identification in sibling models and Section 3.1 for further studies.

on the academic and non-academic development of children that could positively affect outcomes such as the choice of secondary school track. Only recently have curricula at the state level been introduced to specify the competencies children should acquire in preschool (Spieß 2009, p. 378). Effects from German preschool seem to be limited to special groups and, if any, compensate for a lack of very fundamental skills not acquired at home, such as language proficiency of children with migrant background.

Section 2 of this paper provides a brief overview of the German preschool and school system. In Section 3, we present our estimation strategy and describe the GSOEP data. Section 4 sets out both the OLS results and estimates from family fixed effects models on the effect of years of preschool attendance on secondary school track choice, followed by robustness checks. Section 5 concludes.

2 Germany's Preschool and School System

In Germany, kindergarten is the preschool facility attended by children aged three to six. In contrast to the United States, the German kindergarten (hereafter, preschool) is not part of the school system and not integrated with the public schools. German law mandates that preschool is to support parents in reconciling job and family. In addition, preschool is supposed to prepare children for school and (even though not institutionally included with schools) can be regarded as the first part of the education system (Spiess et al. 2003). Attendance is voluntary, but most German children attend preschool for at least one year during the three years previous to school start.[3] Preschools are mostly provided by municipalities and nonprofit organizations; the share of private for-profit providers is negligible.[4] Although preschool attendance is predominantly publicly financed, parents usually have to pay a fee, the amount of which varies across states and/or municipalities. Only recently have state-level regulations with regard to curricula or staff-child ratios been introduced. These regulations are very heterogeneous and not binding for the single child care center (Spieß 2009, p. 378). Thus, attendance at German center-based preschool cannot be compared to participation in programs such as, for example, Head Start, that include health and nutrition services for the children, fixed curricula, or substantial parental involvement.

The secondary school track attended after primary school, which is our dependent variable, has considerable implications for children's future education and occupation. After grade 4 of primary school, the German school system separates students into

[3]Based on own calculations with the GSOEP, 97 percent of the children in our sample attend preschool for at least one year; see also Section 3.2.

[4]In 2008, the share of for-profit providers of preschool (for children aged three to six) was about 0.7 percent in West Germany and about 0.4 percent in East Germany (see Spieß 2009, p. 377).

three different types of secondary schools:[5] The lowest academic track is basic school (Hauptschule) and lasts five or six years. The intermediate track, middle school (Realschule), lasts six years and the highest academic track, the high school (Gymnasium), lasts eight or nine years. The first two tracks are more vocationally oriented; passing the exit exams of German high school (the Abitur) is the traditional and direct way into higher education institutions such as the university.

Different learning and developmental environments in these tracks can lead to differences in cognitive skills at later ages (Baumert et al. 2009). Moreover, as mobility between the different tracks is infrequent (especially upward mobility, see Jürges and Schneider 2007), track choice decisions are highly determinative of the secondary school degree. This, in turn, has implications for individual labor market outcomes such as future earnings (Dustmann 2004).

The decision as to secondary school track is not based on an objective test of cognitive achievement at the end of primary school that could inform about students' aptitude for one of the academic tracks. Instead, at the end of grade 4, primary school teachers make a subjective recommendation for one of the tracks based on students' grade point average in crucial subjects (mostly in German, math, and science). In some German states, parents are not bound by the teacher's recommendation. Lower educated parents who care less about the educational pathways of their children or who are less confident might follow the recommendation or even choose a less prestigious track than recommended. More educated parents will perhaps override the teacher's recommendation and send their children to a higher track than recommended. Even if the teacher recommendation is binding, teachers' (grade-based) assessment does not necessarily reflect students' academic performance. Lüdemann and Schwerdt (2010) show that children with migration background tend to receive recommendations for the less prestigious tracks, even if they have the same performance on student achievement tests and the same general intelligence as their classmates from better-off families.[6] Thus, secondary school track attendance is a multidimensional outcome that is not exclusively affected by differences in cognitive development.

This educational system leads to different interpretations of the channels of a causal association between years of preschool attendance and choice of secondary school track. On the one hand, longer preschool attendance might simply improve the cognitive de-

[5]In the states of the former German Democratic Republic, the lowest academic track, basic school, does not exist and in some other states this track has also been abolished. Moreover, in the states of Brandenburg and Berlin, tracking takes place after grade 6 instead of after grade 4.

[6]Since the recommendation is based on school grades, children with migration background also receive lower grade point averages than their better-off classmates, given equal general intelligence and performance on student achievement tests. Performance on student achievement tests, as well as measures of general intelligence, are from the German extension of the Progress in International Reading Literacy Study 2001 (PIRLS 2001).

velopment that teachers consider crucial for making a recommendation to high school. On the other hand, other less favorable factors that negatively affect parents' track choice and/or teachers' recommendations (such as a lower social background) might be compensated for by skills acquired in preschool. One could expect that preschool enhances the skills that children from a low social and/or migration background need to have over and above those possessed by children from a favorable background in order to either receive a high school recommendation or be sent to one by choice of their parents.[7]

3 Empirical Strategy and Data

This section describes our estimation strategy for identifying causal effects of years of preschool attendance on future school track choice after primary school. Moreover, we provide an overview of the German Socio-Economic Panel (GSOEP), the household survey used in our analysis.

3.1 Estimation Strategy

Our empirical strategy consists of two parts. In the first step, we estimate multivariate OLS regressions analyzing the association between years of preschool attendance and future secondary school track choice. This ties in with several studies that provide first evidence on this question for Germany (Spiess et al. 2003; Büchner and Spiess 2007; Landvoigt et al. 2007; Seyda 2009). In the multivariate OLS analysis, our empirical setting is generally the same as that used in these other papers. However, we have a much larger sample of the GSOEP and we focus on children's preschool attendance during the three years prior to school entrance since this is the period in which children typically attend the German kindergarten.[8] As the dependent variable we choose a dummy indicating whether the child attends the highest school track (that is, high school) after primary school. Formally, this leads to the estimation of the following equation:

$$gym_i = \alpha + \beta * preschool_i + \gamma * cov_i + \epsilon_i \quad (1)$$

[7] In this context, noncognitive skills acquired in German preschools could also play a major role (see, for example, Cinnirella et al. 2011).

[8] Spiess et al. (2003) analyze the effects of preschool attendance in the year before school entrance, whereas Landvoigt et al. (2007) study preschool attendance of children in the five years before school entrance. The latter approach is similar to ours although we cannot estimate the returns to four versus three or five versus four years of preschool attendance. We chose to accept this limitation in the interests of having a larger sample. Büchner and Spiess (2007) even look at six years before school entrance. Seyda (2009), who also analyzes three years before school entrance, comes closest to our approach.

gym_i is a dummy variable that equals 1 if the child attends high school (Gymnasium); 0 otherwise. Our explanatory variable of main interest, $preschool_i$, indicates the time spent in preschool during the years prior to school entrance (from zero to three years). The vector cov_i consists of several variables that could confound the association between years of preschool attendance and high school attendance. We include the highest educational degree of the parents, as well as mother's employment status and household income in the year children enter preschool. Furthermore, we control for a set of children characteristics such as sex, birth order, and the age at school and track entrance. Dummies for the year and state of secondary school track entrance account for time trends and state differences in tracking regimes (Section 3.2 provides a detailed overview of our data set and the variables we use). ϵ_i is a zero-mean error term. The coefficient on $preschool_i$, β, shows the effect of one additional year of preschool attendance on the probability of attending high school.

The GSOEP allows us to include several important covariates in Equation (1) that might affect both our variables $preschool_i$ and gym_i and could therefore bias β in a multivariate setting. However, there are still several concerns over whether the estimated relationship between years of preschool attendance and secondary school track choice is causal in this model. Preschool attendance is not random, but is affected by several factors at the individual and family level. Apart from the characteristics for which we can control in Equation (1), several others are partly or completely unobserved, for example, ability of the children or parental appreciation of education. If these factors affect both preschool attendance and future secondary school track choice, the results from estimations of Equation (1) will be biased.

The direction of a possible bias is not clear a priori. One could expect that the children of parents having a high appreciation of education and a more beneficial social background attend preschool longer than others, and have a higher probability of attending high school even in the absence of preschool attendance. This would lead to an upward bias of our coefficient on $preschool_i$. For example, Becker (2010) finds that German children attend preschool slightly longer than children with a Turkish migration background. On the other hand, some German evidence also shows that children with a lower socio-economic background have a higher probability of attending preschool longer, especially when the availability of preschool is rationed. Fuchs and Peucker (2006) argue that there is a tendency to favor children from less favorable backgrounds who have working mothers when access to preschool for three year old children is limited. This might result in a downward bias of our coefficient of interest.

To more accurately predict whether there is a causal relationship between years of preschool attendance and secondary school track choice, in a second step, we apply a sibling approach. There is a large body of literature dealing with sibling models,

including many discussions of the advantages and pitfalls of this method.[9] The crucial difference between our empirical strategy in Equation (1) and the sibling approach involves including a family fixed effect in the sibling specification. Doing this allows us to account for any unobserved factor at the family level that has the same linear, additive effect on preschool attendance and secondary school track choice of siblings. For example, the role played by parents' appreciation of education, which is constant across their children, can be captured by family fixed effects. The straightforward way of estimating a model with family fixed effects is the following:

$$gym_{if} = \kappa + \delta * preschool_{if} + \tau * cov_{if} + \mu_f + \sigma_{if} \tag{2}$$

The dependent variable gym_{if} indicates whether child i of family f attends high school (Gymnasium) or not; $preschool_{if}$ provides information on the preschool attendance of child i in family f in the three years before school; cov_{if} is a vector of different sibling-specific covariates determining both $preschool_{if}$ and gym_{if}. In this approach, we can include only those factors that differ across siblings, such as household income and mother's employment (both assessed in the year of preschool entrance) or birth order. μ_f is the fixed effect at the family level; σ_{if} is a sibling-specific zero-mean error term. The effect on future secondary school track choice will be identified by within-family variation in the years of preschool attendance across siblings.

This approach excludes many confounding (un)observable factors at the family level that are constant across siblings; however, there are several within-family differences across siblings that might affect their preschool attendance and also our outcome of interest. For example, a family's socio-economic situation could change as the number of children increases. The first born might attract more attention from the parents and they thus may invest more in that child's early education. In such a case, effects of longer preschool attendance of earlier-born children could not be disentangled from stronger parental support of these children in early years. Parents' higher investment in (costly) preschool for earlier-born children could also be due to lower budget constraints when family size is smaller. Therefore, our effects could be confounded with better outcomes of earlier-born children within families, a finding that stems from the

[9] See, for example, Taubman (1976) for one of the first studies using siblings to identify the returns to schooling. A first overview article on this topic is Griliches (1979). A seminal paper is Ashenfelter and Krueger (1994), which uses monozygotic twins to estimate returns to schooling on future earnings. Ashenfelter and Zimmerman (1997) use data on brothers to estimate returns to schooling. More recent work by Garces et al. (2002) uses siblings to identify effects of Head Start attendance on different educational outcomes. Berlinski et al. (2008) exploit differences in sibling-specific preschool attendance in Uruguay in order to study effects on years in education and on drop-out rates. Salm and Schunk (2011) apply a sibling approach to study health as a channel for the transmission of human capital.

literature on birth order effects.[10] Our controls for household income and parents' employment status at the start of preschool, as well as for birth order, are likely to eliminate many of these confounding effects.

Moreover, we account for sibling differences in the age at school entrance since this might be associated with longer preschool attendance and, at the same time, has a direct effect on educational outcomes. In Germany, age at school entrance can vary between six and seven years due to cut-off rules for entrance. Children who start school one year later could have attended preschool one year longer than their siblings. At the same time, the literature on school entrance effects shows a positive association between age at school entrance and future outcomes.[11]

Sibling-specific preschool attendance due to differences in a family's socioeconomic situation over time, birth order, or age at school entrance might well be captured by our controls, but such is not the case for general ability gaps across siblings that are independent of any of the above mentioned factors and that could also explain within-family variation in preschool attendance. Parents might decide about their children's years of preschool attendance with respect to the children's ability. Ability, in turn, is likely to be correlated with high school attendance.[12] For a sub-sample of our children, we have information on innate ability, measuring innate competencies that are not affected by educational pathways and cultural peculiarities. In robustness checks, we show that controlling for innate ability does not change our main result. Likewise, in a separate sample of three year old children, we find that across-siblings variation in social skills is not related to preschool attendance of these children at age three.

Attenuation bias due to random measurement error in our explanatory variable $preschool_{if}$ might be a pitfall to our identification. Measurement error is especially relevant in panel analyses since much of the true signal in the data has already been eliminated by identifying differences across the panel dimension (in our case, across siblings) (see Angrist and Krueger 1999). However, we consider this a minor issue. As we can make use of panel data and track children from the third year prior to school until entrance into secondary school, we do not need to rely on retrospective assess-

[10] See, for example, Black et al. (2005), who find negative effects of birth order on children's education, or Black et al. (2011), showing a negative causal relationship between birth order and the IQ of young men.

[11] For example, Puhani and Weber (2007) find a positive effect of age at school entrance on the probability of attending high school and on student achievement at the end of grade 4 in Germany. Bedard and Dhuey (2006) provide evidence for a positive effect of relative age at school entrance on student performance in TIMSS.

[12] Probably the best example of differences in ability, all else equal, is the case of polyzygotic twins. They share all the factors discussed above, such as birth order, living conditions when growing up, and so on, but can and do differ in ability.

ment of preschool attendance. Moreover, contemporaneous information on preschool is acquired from the persons who finally decide about attendance of the children, that is, the parents. This is a big advantage compared to related studies, which have to use retrospective information about Head Start and/or preschool attendance, such as Garces et al. (2002) or Berlinski et al. (2008). This might increase the risk of recall errors.[13]

3.2 Data on Preschool Attendance and Secondary School Track Choice

The data we use for our analysis are derived from the German Socio-Economic Panel (GSOEP). This representative longitudinal data set has been existing since 1984 and provides annual information about households and all their members on several socio-economically relevant topics.

For our purpose, it is crucial to have information on both preschool attendance and secondary school track attended after primary school. We can exploit that the head of the household provides annual information about all household members younger than 17 years old who are not personally interviewed. The head of the household reports on educational and biographical issues, among them preschool attendance and secondary school track attendance. We use the panel structure of the data set in order to include in our sample children with information on both preschool attendance in the three years before school start and the secondary school track attended after primary school. This requires tracking the children, on average, from age three/four to age ten/eleven, that is, for about seven years.[14]

The head of the household reports whether or not the child is attending preschool at the time of the household interview (usually in the spring). As we do not know exactly what month the child entered preschool, we use a proxy for the years of preschool attendance. For example, if the head of the household reports a child's preschool attendance in every interview in the three years before the child's school start, we assume that this child attended preschool all three years. The other extreme situation is a no-preschool record in all three years before school start. The same procedure is used for a one- or two-year period of preschool attendance. A two-year period of

[13]Garces et al. (2002) have to use retrospective information on Head Start attendance and/or other preschool from adults aged 30 or below. In the sample used by Berlinski et al. (2008), information on years of pre-primary education is provided when children are, on average, 11 years old. Both studies compare their preschool records with numbers from contemporaneous statistics, and show that their preschool data are not likely to be subject to recall error.

[14]For children living in the German states of Berlin and Brandenburg where tracking takes place after grade 6 instead of grade 4, we need an even longer period. Moreover, the length of tracking period will also change in the event a child repeats or skips a grade in primary school.

attendance can mean the child attended preschool in the two adjacent years before primary school, or in the third and second year before start of primary school, or in the third and last year. The same is true for the one-year period.[15]

Our dependent variable is measured in the first year in which the head of the household reports a child's attendance at a secondary school track.[16] This is normally at the age of 10 or 11, depending on the age at entrance into primary school,[17] the tracking regime of the respective federal state, and grade repetition or skipping during primary school.

The sibling sample is smaller since we have to drop all children who do not have at least one sibling with information about preschool attendance in the three years before primary school and secondary school track attendance after primary school. This will be the case if children do not have any siblings or if we cannot observe all necessary information about an existent sibling in our data set.[18] We provide more information on the sibling sample when presenting our sibling models estimations in Section 4.2.

Table 1 provides an overview of the descriptive statistics of the full sample. As the first interview in the GSOEP took place in 1984, in our sample the oldest children started school in 1987, allowing us to observe the full three-year period before entrance. In total, we include children born between 1979 and 2000. The youngest children just enter secondary school (and thus provide information on secondary school track) in the most recent wave 2009. The variable *Years of preschool attendance* shows the distribution of the time spent in preschool, our main variable of interest: 73 percent of all children in our sample attend preschool all three years before school entrance. Only about 3 percent of the children never attend preschool. This corroborates findings from official statistics showing that almost all children attend preschool for at least one year before school start.

[15]Only 17 percent of all children who go to preschool two out of three years before school start do not attend in the two adjacent years to school start. They attend in the second and third year or in the third and last year before school start. Likewise, only about 29 percent of all children who go to preschool one out of three years before school start do not attend in the year before school start but in the second or third year prior to school. We include these children in the sample to maximize the number of observations.

[16]Children attending comprehensive school (Gesamtschule) after primary school are also included in the sample. Table A1 shows estimates from a sample without these children. The results are very similar to those in Table 5.

[17]Age at entrance into primary school partly depends on the child's month of birth since Germany uses cut-off rules for school entrance. As there are exceptions to this rule, redshirting and earlier entrance than prescribed by the cut-off rules can also affect the age at entrance.

[18]This situation is more probable at the beginning (1984) and the end (2009) of our sample: we do not have complete preschool information for those children attending preschool before 1984. Similarly, we cannot observe track attendance for children entering secondary school after 2009 and drop them from our sample.

The average age at school entrance and track entrance cannot be precisely determined in the GSOEP, due to the timing of the interviews, which mostly take place in spring and thus, on average, half a year after the start of the school year in August or September (depending on the German federal state). Fortunately, we can link our household data to exact information about the month of school start (August or September) in every German federal state for the entire sample period. This allows us to identify the age at school entrance and secondary school track entrance in months for each child. The information on the month of school start for each German state between 1984 and 2009 is derived from historic holiday calendars available on the website of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (see Kultusministerkonferenz 2010).[19] On average, children are 80.1 months old at the start of primary school and 128.2 months old when entering secondary school track. More than a third of all children in our sample (35 percent) attend high school after primary school. The categorical variable employment status of the mother, measured in the year in which the child enters preschool,[20] shows that more than half the mothers (55 percent) do not work in the year in which their children enter preschool. The highest educational degree of fathers and mothers does not differ substantially. Most of the parents finished either basic school (Hauptschule) or middle school (Realschule).

All variables provide an average across children born over more than two decades (1979-2000). This is especially important with regard to the secondary school track attended after primary school since the patterns have changed over time: until the mid 1990s, the share of children who went to high school was lower than our figure of 35 percent.[21] It is therefore crucial to include dummies for the year of entrance to secondary school in all our specifications to capture overall time trends in track choice.

4 Results

The sub-sections below provide estimation results of the relationship between years of preschool attendance and the probability of attending high school (Gymnasium). First, we show basic OLS estimates, followed by our sibling specifications. After providing

[19] For almost all children, we know the age in months at the interview and therefore can calculate their age in months at the last possible school/track entrance. For those few children without information on the month of interview and/or month of birth, we compute age at school and secondary school track entrance using the difference between the year in which they first report school and secondary school track entrance and their year of birth.

[20] For those children who never attend preschool, this variable is measured in the year before school entrance.

[21] Among the children who entered secondary school up to 1995, only about 27 percent attended high school (Gymnasium), whereas for those who entered after 1995, this share is about 37 percent.

several robustness checks, we end this results section with findings from using cognitive skills as an alternative outcome.

4.1 OLS Results

As a benchmark, we first present results from multivariate linear probability and probit estimations on the association between years of preschool attendance and the probability of attending high school. Column (1) of Table 2 shows a significant positive bivariate association between years of preschool attendance and the probability of attending high school from a linear probability model. This result is corroborated in the probit estimation in Column (2). When we include additional control variables at the child and family level (Columns (3) and (4)), as well as dummies for year and state of secondary school track entrance, the point estimates on years of preschool attendance become smaller, but remain positive and significant.[22] In Columns (5) and (6) we limit our sample to children with parents who have less than an upper secondary degree (Column (5)) and less than a degree from middle school (Column (6)). The coefficient estimates on years of preschool attendance are even larger than in the linear full sample specification of Column (3), but not significantly different from the estimate in Column (3).

Preschool is known to be particularly beneficial for disadvantaged children and/or children with a migration background: this result is found in a great deal of the early childhood education literature, mostly based on evaluations of Head Start and model programs (Blau and Currie 2006, p. 1234). In Table 3 we therefore focus on the sub-sample of children who have, as defined by the German Federal Statistical Office, a migration background.[23] We estimate the same linear probability models as in Table 2 and find significant positive estimates on years of preschool attendance. The point estimates in Columns (2)-(4) are all larger in magnitude than in the respective specifications in Table 2. Although the difference is not statistically significant, this result hints at a stronger association between years of preschool attendance and the probability of attending high school for children with a migration background.

The results from Table 2 and Table 3 mainly confirm what has been found in other German studies, such as Büchner and Spiess (2007) and Seyda (2009), but their causal interpretation remains difficult. In the following sub-sections, we tackle this problem by reporting results from sibling models.

[22] From now on, we report results only from linear probability models as probit results are very similar.

[23] According to the German Federal Statistical Office, people have a migration background if they either immigrated to Germany after 1949, were born in Germany as foreigners, or they are German citizens but have at least one parent who immigrated to Germany after 1949. In our sample, 98.5 percent of the migrant children are second-generation immigrants.

4.2 Sibling Models

In the sibling models we identify the effect on high school attendance exploiting within-family variation in years of preschool attendance across siblings. Table 4 provides an overview of the sibling sample and its most important characteristics.

Most of the children who have siblings have one sister or brother (more than 70 percent of the children). The average age difference between the youngest and oldest child of a family is about 4.1 years in the full sibling sample and about 4.6 years in the migration background sample. In most families who have more than one child in our sample, we observe no difference in the years of preschool attendance across siblings. The preschool effect in the family fixed effect specifications is identified by children from 193 families (73 in the sample of children with migration background). When preschool attendance does vary within families, it mostly differs between siblings by one year. The share of children living in families with two-year differences in attendance is higher among children with a migration background (21.69 percent vs. 11.82 percent). In families where children have spent different amounts of time in preschool, more than half of the longer-attending children are the second born. However, there are also first-born children who attend preschool longer than their younger siblings. This allows us to control for birth order in our family fixed effects specifications. Twenty-seven percent (17 percent in the migration background sample) of the children from families in which siblings differ in years of preschool attendance also live in families with variation in high school attendance across siblings.

In Table 5 we compare results from specifications with and without family fixed effects using the sibling sample. In Panel A we reestimate the linear probability models from Tables 2 and 3 without including family fixed effects. The point estimates on years of preschool attendance are positive and significant across all specifications. An additional year of preschool attendance is associated with a higher probability of attending high school ranging between 3 and 9 percentage points (depending on the sample). In both the full sibling sample and the sibling sample of children with migration background the point estimates gradually increase in magnitude when we limit the sample to children with lower educated parents. Again, the estimates in the sibling sample of children with migration background (Columns (5)-(7)) are higher in magnitude than the corresponding estimates in the full sample (Columns (2)-(4)). More importantly, none of the point estimates in Panel A of Table 5 is statistically distinguishable from the respective coefficient estimates in Tables 2 and 3. Stable associations between years of preschool attendance and high school attendance indicate that our sibling sub-sample is probably not selective with regard to the full sample.

Panel B reports our main results. It shows estimates from the same specifications as in Panel A, but now supplemented by a family fixed effect. In Column (1), which shows estimates from the bivariate association between years of preschool attendance and high school attendance, we find a non-significant point estimate of -0.033. This estimate is significantly smaller than the corresponding coefficient estimate from Panel A of 0.071. Including all sibling-specific control variables and restricting the sample to children with lower educated parents (in Columns (2)-(4)) does not change the general pattern. The coefficient estimates in the family fixed effects specifications are significantly smaller than the corresponding estimates that do not include fixed effects and they are not significantly different from zero. The migration background sample (Columns (5)-(7)) shows similar results: although not always statistically distinguishable from the coefficient estimates in Panel A, the points estimates on years of preschool attendance become much smaller in the family fixed effects specifications of Panel B. The coefficients on years of preschool attendance in the sibling sample of children with migration background are (in contrast to the full sibling sample) positive but, again, not statistically different from zero.[24]

We showed that siblings vary in number of years of preschool attended in only 193 out of the 540 families in our sample having more than one child. We therefore test whether the family fixed effects results are robust to the exclusion of families without within-family variation in years of preschool attendance. Panel C of Table 5 shows that family fixed effects estimates do not significantly differ from the corresponding coefficients in Panel B if we drop families in which years of preschool attendance do not vary across siblings.

Years of preschool attendance might not positively affect the probability of attending high school, but could increase the probability of attending a higher secondary school track than basic school (Hauptschule), which is the lowest track. Table 6 shows that this is not the case. In most specifications, we observe similar differences in point estimates between the specifications with and without family fixed effects as in Table 5. Most importantly, there is no significant effect of years of preschool attendance in any of the family fixed effect specifications.

Effect Heterogeneity

As shown in Table 4, within-family variation in preschool attendance is mostly one year. One might argue that an additional year of preschool does not increase the probability of attending high school, but perhaps two or three additional years would. To investigate effect heterogeneity, we built a sample of siblings who differ by at least

[24]We found the coefficients in Panel A of Table 5 to be significantly distinguishable from the corresponding coefficients in Panel B in Columns (1), (2), (3), (4) and (7).

two years of preschool attendance (in Table 7) and reestimate the specifications from Table 5. We now find positive point estimates in the full sibling sample specifications of Columns (1)-(3) in Panel B. In the sibling sample of children with migration background, coefficients on years of preschool attendance hardly differ from those of Table 5. However, all coefficients on years of preschool attendance in the family fixed effects specifications remain insignificant and smaller in magnitude compared to the estimates without family fixed effects.[25] We conclude that even children who attend preschool two or three years longer than their siblings do not have a significantly higher probability of attending high school.

We further check whether there are heterogeneous effects with respect to daily exposure to preschool. The GSOEP provides information on whether children attend preschool half-day or full-day.[26] In Table 8 we interact the years of full-day preschool attendance (varying between zero and three years) in the three years prior to school with our main variable of interest, that is, years of preschool attendance. This does not change our basic estimates from Table 5. Focusing on the results from Panel B of Table 8, we do not find a significant interaction effect in any of our specifications. Neither is it clear whether the effect of an additional year of full-day preschool attendance (the sum of the coefficients on *Years of preschool attendance* and *Interaction with full-day attendance*) is smaller or larger than an additional year of half-day attendance which is the coefficient on *Years of preschool attendance*. Most importantly, family fixed effects estimates on both one additional year of half-day preschool and one additional year of full-day preschool remain insignificant and also smaller than in the corresponding specifications without family fixed effects from Panel A.

To this point, within-family variation in years of preschool attendance does not produce significant effects on high school attendance. The family fixed effects estimates are smaller than the corresponding coefficients from the specifications without family fixed effects. At first sight, this is especially surprising with regard to children with a migration background. For example, recent German evidence has shown that longer preschool attendance is positively associated with German-language skills of Turkish children (Becker 2010). Thus, preschool attendance might compensate for insufficient conversation in German at home and increase the probability of high school attendance for children from non-German-speaking families. In Table 9 we focus on the sample of children with migration background and interact our variable years of

[25] Significance tests show that the coefficients in Panel B of Table 7 are significantly distinguishable from the corresponding coefficients in Panel A in Columns (2), (3), and (6).

[26] As shown in Table 1, among the children who attend preschool for at least one year in the three years prior to school, 61.2 percent never attend full-day, 15.1 percent report one year of full-day preschool attendance, and 11.4 percent report two years of full-day attendance; 12.3 percent are exposed to full-day preschool in all three years before school start.

preschool attendance with a dummy variable indicating whether the child's mother speaks the family's language of origin (dummy=1) or whether she speaks mostly German (dummy=0).[27] The significant positive interaction effect shows that the impact of an additional year of preschool attendance is significantly larger for children with a migration background and mothers who do not speak German than for those children with a migration background but German-speaking mothers (Column (1)). If we include control variables and limit our sample to children with less educated parents, the point estimates on the interaction effect become insignificant; however, they remain similar in magnitude to the estimate from Column (1). We interpret this finding as a modest hint at preschools' capacity to improve language skills of children with migration background.

However, the effects of years of preschool attendance on the probability of attending high school for the children with non-German speaking mothers (the sum of the coefficient on *Years of preschool attendance* × *Mother speaks no German* and on *Years of preschool attendance*) are also not significantly positive in the specifications of Table 9. The effects on the children with migration background and German-speaking mothers in Columns (2)-(4) of Table 9 (the coefficient on *Years of preschool attendance*) are very similar to the non-significant negative point estimates found in Panel B of Table 5 in the full sibling sample. Effects of preschool for children with and without migration background obviously do not differ after variation in language proficiency is captured.

In unreported specifications we interacted years of preschool attendance with several variables such as birth order, sex of the children and mother's employment status. In none of these estimations we find any substantial differences from our general pattern. Coefficients on the interaction terms are close to zero and the estimates in family fixed effects models remain insignificant and considerably smaller than in the models without fixed effects.

In sum, including family fixed effects in sibling samples countervails significant positive associations between years of preschool attendance and the probability of attending high school found in multivariate estimations without family fixed effects. This is even true in sub-samples of children with migration background, if we only look at two and three years' difference in preschool attendance across siblings and if we consider daily exposure to preschool. We find only modest evidence for preschools' capacity to enhance German-language skills for children with migration background who have

[27]The corresponding question comes from the SOEP person interviews. All persons with a migration background are asked about the language they speak in Germany. There are three answer categories: *Mostly German*; *mostly language of origin*; *half German, half language of origin*. We build a dummy variable that equals 1 if the mother reports speaking *mostly language of origin* or *half German, half language of origin* and 0 if the mother reports *Mostly German*. We drop those children from the sample whose mothers did not answer this question.

non-German speaking mothers. In the next sub-section, we show that our results are robust to other sibling-specific differences in several dimensions that might affect within-family variation in years of preschool and high school attendance. Furthermore, we provide first estimation results from using cognitive skills as an alternative outcome of years of preschool attendance.

4.3 Robustness Checks

In Section 3.1, we discussed possible sources of within-family differences in preschool attendance across siblings. Several of the potentially confounding factors mentioned in that section were captured by including sibling-specific covariates, but ability differences across siblings have not yet been addressed. Identification in the family fixed effects models could be hampered if ability differences have an impact on the probability of attending high school. In our case, for example, the small estimates in family fixed effects specifications could be downward biased if parents send their less able children to preschool longer.

Innate Ability

The GSOEP allows us to study whether within-family differences in innate ability and social skills across siblings affect preschool attendance. Panel A of Table 10 shows descriptive statistics from performance on a cognitive ability test given to 17 year old adolescents who are included in our original sample (that is, there is information on these adolescents' preschool attendance during the three years prior to school start and on their secondary school track attendance). The test consists of three parts assessing numerical, verbal, and figural skills. We use only the test on figural skills since is meant to be independent from educational pathways and cultural peculiarities and credibly measures innate, genetically determined ability. This is important as we want to capture ability differences across siblings that have not been affected by preschool or later schooling. Our outcome of interest is the number of correct answers in the figural skills test (on a scale from 0-20). We observe only small differences in mean performance and performance distribution between the sample of all children taking the test and the sample in which we observe test performance of at least two siblings. However, we see that within-family variation in innate ability across siblings is smaller than across two random children of the sibling sample (2.06 vs. 3.16). While this is not surprising, we also observe that innate ability differences at the family level are even smaller if the siblings differ in preschool attendance than if they do not. This suggests that innate ability differences are probably not a channel of within-family variation in preschool attendance.

Further, we include the measure of innate ability from the assessment of figural skills in our family fixed effects models that estimate the association between years of preschool attendance and high school attendance. This is possible for only the very small sibling sub-sample of 17 year old adolescents who participated in the assessment (carried out in the GSOEP waves 2006-2009) and who provide information on preschool attendance in the three years before school and on secondary school track choice.

Columns (1) and (2) of Table 11 present baseline results using linear probability models without including family fixed effects. In Column (2), we include our measure of innate ability. The coefficient estimates on years of preschool attendance do not significantly differ in Columns (1) and (2): both are positive and insignificant, but similar in magnitude to the point estimates in Panel A of Table 5 for the full sibling sample. As suggested, we see a significant positive correlation between innate ability and the probability of attending high school in Column (2). Turning to the family fixed effects specifications (in Columns (3) and (4)), we observe a similar pattern: if we include innate ability in Column (4), the point estimate on years of preschool attendance remains insignificantly negative, as in Column (3), and only slightly increases in magnitude. The point estimate on innate ability is insignificant in Column (4) of Table 11.[28] This corroborates the results that we found in Panel A of Table 10 that differences in innate ability across siblings do not significantly affect within-family differences in preschool attendance. Given the same innate ability across siblings, one additional year of preschool attendance does not significantly increase the probability of attending high school.

Social Skills

In Panel B of Table 10 we report similar descriptive statistics as in Panel A for a sample of three year old children for whom we have information on social skills and preschool attendance (reported by the mother). Since first interviews with the mothers took place in 2002, these children are not part of our original sample. We cannot fully exclude that the children's social skills at age three have already been affected by educational experiences before preschool. Yet, attendance rates at out-of-home care before the age of three are very low and thus we are confident that we are using a measure that largely captures innate social skills.[29] Mothers assess their children's social skills by responding to the following five statements: (I) child calls familiar people

[28] As observed across all other specifications, the point estimates on years of preschool attendance in family fixed effects models are smaller than the corresponding point estimates without family fixed effects.

[29] In 2006, the year of birth of the youngest three year old children in our sample, 13.6 percent of children younger than three years attended out-of home care (Statistische Ämter des Bundes und der Länder 2007).

by name, (II) child plays with other children, (III) child participates in role-playing games, (IV) child has preferences for certain friends, and (V) child calls own feelings by name. Mothers can answer *yes*, *to some extent*, or *no*. The number of *yes* answers (1-5) is our social skills outcome. Preschool attendance is a dummy variable indicating whether or not the child attends preschool. We observe the same patterns as in Panel A of Table 10. Within-family differences in social skills do not vary across families in which we observe variation in preschool attendance and those in which preschool attendance does not differ across siblings. The within-family standard deviation is 0.70 for both family types.

We next estimate the relationship between children's social skills and their preschool attendance using the sample of three year old children. Since these children are not part of our original sample, we cannot include children's social skills as a further control variable in estimations of years of preschool attendance on high school attendance. As indicated, the mothers report only on whether or not the child is attending preschool at the time of the interview. In our original sample, we observe preschool attendance in the three years before school and show that most within-family variation in preschool attendance across siblings is between two and three years. So, if social skills differences really affect parents' decision about their children's preschool attendance, this decision will be made when children are three years old. Our sample of three years olds is well suited to test this.

Table 12 shows estimates from linear probability models. The dependent variable, preschool attendance, is coded as a dummy variable that equals 1 if the child attends preschool; 0 otherwise. Children's age in years is three; however, their age in months varies between 36 and 47 months. We want to capture these age differences and include children's age (in months) as a further explanatory variable. Column (1) of Table 12 presents a highly significant positive association between children's social skills and their preschool attendance based on the whole sample. This is robust to inclusion of the child's age (in months) as a further control variable (in Column (2)). Likewise, the coefficient estimate on social skills remains positive and statistically significant if we include in the sample only children with at least one sibling (Column (3)). When we include family fixed effects in Column (4), the association between social skills and preschool attendance becomes smaller and insignificant. In sum, the results hint at a strong correlation between social skills and preschool attendance, but not across siblings within families. Children who attend preschool at age three do not significantly differ in their social skills from their three year old siblings who do not attend.[30]

[30]The effect size that we cannot exclude to be zero in Column (4) of Table 12 is quite small. An increase of the social skills by one unit (which is about 1.3 within-family standard deviations) raises

We are confident that we have accounted for most of the possibly confounding determinants affecting within-family preschool attendance and secondary school track choice. This naturally raises the question about the remaining identifying variation in preschool attendance across siblings. Of course, some differences in siblings' preschool attendance are due to an expansion of public preschool over time. However, this cannot be the only reason because we show that in some families the oldest children are those with the longest exposure to preschool. Other sources of variation might be local supply differences that we cannot capture with our data. Further, if mothers have to stay at home with their new born offspring, they may also decide to provide at-home care for their older children of preschool age. This could reduce the years of preschool attendance of lower-rank children. However, this explanation is again not consistent with longer preschool attendance by the oldest children in some families. In sum, we cannot definitely identify the source of residual differences in within-family preschool attendance. However, and importantly, as shown by our robustness checks, potential reasons are most likely not correlated with our outcome of interest.

4.4 Cognitive Skills as Dependent Variable

Non-significant effects of years of preschool attendance on high school attendance might be due to the fact that the determinants of secondary school track choice are complex (see Section 2). In particular, cognitive skills acquired in preschool and school is not the only factor affecting high school attendance. Therefore, effects of preschool attendance on cognitive skills may not necessarily be reflected in a significantly higher secondary school track attendance. We study cognitive skills as an alternative outcome possibly affected by years of preschool attendance. While the test on adolescents' figural skills reasonably assesses innate ability, the numerical and verbal sections of the test capture cognitive skills accumulated over the lifetime. This allows us to test the association between years of preschool attendance and cognitive skills, at least in the very small sample of adolescents with information on all relevant variables (years of preschool attendance and cognitive skills).

In Table 13 we use the correct answers from the numerical and verbal sections of the test as our outcome variable. On average, adolescents answer 22 out of 40 (20 numerical and 20 verbal) questions correctly; the standard deviation is 7. Columns (1) and (2) report results from linear probability models without including family fixed effects. We find no significant effects of years of preschool attendance on our cognitive skills measure. The coefficients are negative and close to zero. The same is true for

the probability of attending preschool by about 4 percentage points (with an average probability to attend preschool in the sibling sample of 0.41).

the models with family fixed effects (Columns (3) and (4)). One additional year of preschool attendance has no significant effect on the cognitive skills of 17 year old adolescents.[31]

The use of innate ability as an additional explanatory variable and cognitive skills as an alternative outcome variable is only possible for very small samples, providing quite imprecise estimates. Similarly, robustness checks of the effects of social skills on within-family preschool decisions are based on information about children who are not part of our original sample. The reported results should thus be interpreted with caution. Nonetheless, our main pattern from estimations based on the full sibling sample is stable across all robustness checks. It shows that coefficients in family fixed effects estimations are considerably smaller than in the specifications that do not account for unobserved heterogeneity at the family level.

5 Conclusion

In this paper, we estimated the causal effect of years of preschool attendance on the probability of attending the highest secondary school track in Germany, that is, high school. Attending high school leads to obtaining Germany's highest secondary school degree (the Abitur), facilitates access to tertiary education, and therefore affects individual earnings. German preschool is not comparable to highly intensive, targeted early childhood education programs that exhibit considerable, positive long-term effects. Yet, several multivariate studies show that German center-based preschool is positively associated with secondary school track choice and therefore might have the potential to affect future educational outcomes. We complement these previous analyses by studying whether this association continues to hold when employing sibling models that capture unobserved family-level heterogeneity.

Estimating multivariate models across families, we confirm a positive association between one additional year of preschool attendance and the probability of attending high school. This result is robust across specifications in which we limit the sample to children with lower educated parents or with a migration background.

When we include family fixed effects in sibling models, however, the positive association between years of preschool attendance and high school attendance vanishes. None of our estimates is significantly different from zero and in almost all specifications it is considerably smaller in magnitude compared to the coefficients in models

[31] The non-rejectable effect sizes are all very small, except from the point estimate in Column (3). In the latter we exclude an effect of years of preschool attendance on higher cognitive skills in the magnitude of about a sixth of a standard deviation.

without family fixed effects. This indicates that within-family variation in preschool attendance does not affect the probability of high school attendance.

In the sibling approach, we can capture (un)observed factors at the family level that affect siblings' preschool and high school attendance alike. While this already supports a causal interpretation of our results, we show in robustness checks that certain problems with our identifying assumptions likely pose only a minor threat. Sibling-specific socio-economic covariates such as birth order, household income and employment status of the mother at the start of preschool, as well as innate ability and social skills differences across siblings do not affect within-family preschool decisions and thus do not change our estimates.

We conclude that small, non-significant estimates of years of preschool attendance on high school attendance from family fixed effects models reveal the bias in multivariate models due to unobserved factors determining preschool decisions. Our results hint at a longer preschool attendance by children from more favorable backgrounds and a higher probability of attending high school by these children. The small causal estimates from family fixed effects models probably result from the low-intensity, large-scale nature of German preschool. The literature that finds limited effects of the expansion of kindergarten and prekindergarten in the United States argues similarly (Cascio 2010; Dhuey 2011).[32]

The most we can say is that benefits from an additional year of preschool are significantly larger for children with a migrant background who have non-German speaking mothers compared to those with German-speaking mothers. Discovering whether it is language proficiency that is enhanced by preschool or whether other (cultural) differences between migrant children with German- and non-German-speaking mothers cause these effects must be left to further research.

The low intensity of German preschool might be one reason for our findings, but our choice of outcome variable could also be driving our results. As outlined, secondary school track choice is a complex decision made by several decision-makers such as teachers and parents. Possibly, preschool only affects single determinants of high school attendance, such as cognitive or social skills. But we do also not find positive effects of years of preschool attendance on cognitive skills of 17 year old adolescents. However, our results are based on a very small sibling sample and outcomes are assessed quite some time after transition to secondary school.[33] In any case, potential medium-

[32]In the United States, low intensity is only one reason for modest effects. Small benefits from U.S. kindergarten and prekindergarten, especially for African Americans, might be driven by reduced enrollment in other more intensive programs (such as Head Start) as a consequence of the expansion of publicly funded kindergarten (see Cascio 2010).

[33]Fitzpatrick (2008), for example, shows positive effects of the availability of prekindergarten on math test scores in grade 4, which is exactly the transition point to secondary school in Germany.

and long-term benefits from German preschool should be analyzed in more detail to discover which outcomes are actually affected and which are not. Such analyses will help policy-makers institute suitable reforms and design appropriate interventions with regard to early childhood education.

References

Angrist, J. D. and Krueger, A. B. (1999). Empirical strategies in labor economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3A: 1277-1366. Amsterdam: North Holland.

Ashenfelter, O. and Krueger, A. B. (1994). Estimates of the economic returns to schooling from a new sample of twins. *American Economic Review*, 84(5):1157–1173.

Ashenfelter, O. and Zimmerman, D. J. (1997). Estimates of the returns to schooling from sibling data: Fathers, sons, and brothers. *Review of Economics and Statistics*, 79(1):1–9.

Baumert, J., Becker, M., Neumann, M., and Nikolova, R. (2009). Frühübergang in ein grundständiges Gymnasium - Übergang in ein privilegiertes Entwicklungsmilieu? Ein Vergleich von Regressionsanalyse und Propensity Score Matching. *Zeitschrift für Erziehungswissenschaft*, 12(2):189–215.

Becker, B. (2010). Wer profitiert mehr vom Kindergarten? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62(1):139–163.

Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics*, 121(4):1437–1472.

Belfield, C. R., Nores, M., Barnett, S., and Schweinhart, L. (2006). The High/Scope Perry Preschool Program - cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources*, 41(1):162–190.

Berlinski, S., Galiani, S., and Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, 93(1-2):219–234.

Berlinski, S., Galiani, S., and Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, 92(5-6):1416–1440.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics*, 120(2):669–700.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011). Older and wiser? Birth order and IQ of young men. *CESifo Economic Studies*, 57(1):103–120.

Blau, D. M. and Currie, J. (2006). Preschool, day care, and after school care: Who's minding the kids? In Hanushek, E. and Welch, F., editors, *Handbook of the Economics of Education*, volume 2: 1163-1278. Amsterdam: North Holland.

Büchner, C. and Spiess, C. K. (2007). Die Dauer vorschulischer Betreuungs- und Bildungserfahrungen : Ergebnisse auf der Basis von Paneldaten. SOEPpapers on Multidisciplinary Panel Data Research 10.

Cascio, E. U. (2009). Do investments in universal early education pay off? Long-term effects of introducing kindergartens into public schools. NBER Working Paper 14951.

Cascio, E. U. (2010). What happened when kindergarten went universal? *Education Next*, 10(2):63–69.

Cinnirella, F., Piopiunik, M., and Winter, J. (2011). Why does height matter for educational attainment? Evidence from German pre-teen children. *Economics and Human Biology*, forthcoming.

Cunha, F. and Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, 97(2):31–47.

Dhuey, E. (2011). Who benefits from kindergarten? Evidence from the introduction of state subsidization. *Educational Evaluation and Policy Analysis*, 33(1):3–22.

Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, 56(2):209–230.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The B.E. Journal of Economic Analysis & Policy*, 8(1):1–38.

Fuchs, K. and Peucker, C. (2006). ...Und raus bist du - Welche Kinder besuchen nicht den Kindergarten und warum? In Bien, W., Rauschenbach, T., and Riedel, B., editors, *Wer betreut Deutschlands Kinder?* 63-82. Weinheim und Basel: Beltz Verlag.

Garces, E., Thomas, D., and Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, 92(4):999–1012.

Gormley, W. T. J., Gayer, T., Phillips, D., and Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology*, 41(6):872–884.

Gould, E. D., Lavy, V., and Paserman, M. D. (2011). Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes. *Review of Economic Studies*, 78(3):938–973.

Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy*, 87(5):S37–S64.

Havnes, T. and Mogstad, M. (2011). No child left behind. Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the High/Scope Perry Preschool Program. *Journal of Public Economics*, 94(1-2):114–128.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Jürges, H. and Schneider, K. (2007). What can go wrong will go wrong: Birthday Effects and early tracking in the German school system. CESifo Working Paper 2055.

Kultusministerkonferenz (2010). Website of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, accessed: 04 April 2011. http://www.kmk.org/ferienkalender.html.

Landvoigt, T., Muehler, G., and Pfeiffer, F. (2007). Duration and intensity of kindergarten attendance and secondary school track choice. ZEW Discussion Paper 07-051.

Loeb, S., Bridges, M., Bassok, D., Fuller, B., and Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, 26(1):52–66.

Lüdemann, E. and Schwerdt, G. (2010). Migration background and educational tracking: Is there a double disadvantage for second-generation immigrants? CESifo Working Paper 3256.

Ludwig, J. and Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics*, 122(1):159–208.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–1026.

Magnuson, K. A., Ruhm, C., and Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, 26(1):33–51.

Puhani, P. A. and Weber, A. M. (2007). Does the early bird catch the worm? Instrumental variable estimates of educational effects of age of school entry in Germany. *Empirical Economics*, 32(2-3):359–386.

Salm, M. and Schunk, D. (2011). The relationship between child health, developmental gaps, and parental education: Evidence from administrative data. *Journal of the European Economic Association*, forthcoming.

Seyda, S. (2009). Kindergartenbesuch und späterer Bildungserfolg - Eine bildungsökonomische Analyse anhand des Sozio-ökonomischen Panels. *Zeitschrift für Erziehungswissenschaft*, 12(2):233–251.

Spieß, C. K. (2009). Achillesferse "frühkindliche Bildung". *Wirtschaftsdienst*, 89(6):376–379.

Spiess, C. K., Büchel, F., and Wagner, G. G. (2003). Children's school placement in Germany: Does kindergarten attendance matter? *Early Childhood Research Quarterly*, 18(2):255–270.

Statistische Ämter des Bundes und der Länder (2007). *Kindertagesbetreuung Regional 2006 - Ein Vergleich aller 439 Kreise in Deutschland*. Wiesbaden: Statistisches Bundesamt.

Taubman, P. (1976). The determinants of earnings: Genetics, family, and other environments; a study of white male twins. *American Economic Review*, 66(5):858–870.

Temple, J. A. and Reynolds, A. J. (2007). Benefits and costs of investments in preschool education: Evidence from the child-parent centers and related programs. *Economics of Education Review*, 26(1):126–144.

Weinert, S., Asendorpf, J. B., Beelmann, A., Doil, H., Frevert, S., and Lohaus, A. (2007). *Expertise zur Erfassung von psychologischen Personenmerkmalen bei Kindern im Alter von fünf Jahren im Rahmen des SOEP*. Data Documentation 20. DIW Berlin.

Table 1
Descriptive Statistics: Full Sample

Variable	Mean	Std.Dev.
Years of preschool attendance		
0	0.028	
1	0.038	
2	0.202	
3	0.732	
Years of full-day preschool attendance		
0	0.612	
1	0.151	
2	0.114	
3	0.123	
High school attendance (Gymnasium)	0.35	
Child characteristics		
Sex (male=1)	0.50	
Birth order		
Oldest	0.724	
Second oldest	0.228	
Third oldest	0.043	
Fourth oldest	0.005	
Age at school entrance (in months)	80.08	5.47
Age at track entrance (in months)	128.21	10.29
Family background characteristics		
Father's education		
Basic school	0.347	
Middle school	0.234	
Upper secondary technical school degree	0.077	
High school	0.152	
Other school degree	0.124	
No school degree	0.032	
In education	0.003	
Education missing	0.032	

Table continues

Variable	**Mean**	**Std.Dev.**
Mother's education		
Basic school	0.296	
Middle school	0.380	
Upper secondary technical school degree	0.033	
High school	0.137	
Other school degree	0.110	
No school degree	0.036	
In education	0.000	
Education missing	0.008	
Employment status of the mother (at the start of preschool)		
Full-time	0.138	
Part-time	0.217	
In education and/or training	0.006	
Marginally employed	0.076	
Not employed	0.548	
Employment missing	0.015	
Monthly household income (at the start of preschool, in €)	2,283.98	1,099.50
Number of observations	2321	

The full sample consists of all children providing information on preschool attendance during the three years prior to school start and information on the secondary school track attended after primary school. Std. Dev.: Standard deviations are reported only for continuous and discrete variables. Missing values in all categorical family background variables (*Fathers' education*, *Mothers's education* as well as *Mothers' employment status (at start of preschool)*) have been imputed with a missing category dummy. A few cases missing values in the household income variable have been dropped from the sample. Full-day preschool attendance is reported for all children who attend preschool at least for one year in the three years prior to school.

Table 2
Linear Probability and Probit Models

	Full sample				Parents < upper secondary	Parents < middle school
	Linear	Probit	Linear	Probit	Linear	Linear
	(1)	(2)	(3)	(4)	(5)	(6)
Years of preschool attendance	0.076***	0.085***	0.033***	0.050***	0.051***	0.059***
	(0.013)	(0.016)	(0.012)	(0.017)	(0.013)	(0.014)
Highest educational degree of parents			Yes	Yes	Yes	Yes
Employment status of the mother			Yes	Yes	Yes	Yes
Household income			Yes	Yes	Yes	Yes
Sex			Yes	Yes	Yes	Yes
Birth order			Yes	Yes	Yes	Yes
Age at school entrance			Yes	Yes	Yes	Yes
Age at track entrance			Yes	Yes	Yes	Yes
Dummies for year of secondary school track entrance			Yes	Yes	Yes	Yes
Dummies for federal state of secondary school track entrance			Yes	Yes	Yes	Yes
Observations	2,321	2,321	2,321	2,321	1,629	848

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). Columns (1), (3), (5), and (6) show results from linear probability models; Columns (2) and (4) show marginal effects from probit estimations. *parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 3
Linear Probability Models: Children with Migration Background

	Full sample		Parents < upper secondary	Parents < middle school
	(1)	(2)	(3)	(4)
Years of preschool attendance	0.063***	0.049***	0.065***	0.068***
	(0.017)	(0.017)	(0.018)	(0.016)
Highest educational degree of parents		Yes	Yes	Yes
Employment status of the mother		Yes	Yes	Yes
Household income		Yes	Yes	Yes
Sex		Yes	Yes	Yes
Birth order		Yes	Yes	Yes
Age at school entrance		Yes	Yes	Yes
Age at track entrance		Yes	Yes	Yes
Dummies for year of secondary school track entrance		Yes	Yes	Yes
Dummies for federal state of secondary school track entrance		Yes	Yes	Yes
Observations	681	681	590	500

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns show results from linear probability models. *parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 4
Descriptive Statistics: Sibling Sample

	Full sibling sample	Siblings with migration background
Number of siblings:		
1	72.97%	72.33%
2	23.35%	24.38%
3	3.68%	3.29%
Average age difference between oldest and youngest child in family (in years)	4.07	4.62
Families with more than one child	540	167
Families with more than one child and variation in preschool	193	73
Share of children by within-family difference in preschool attendance (in years) preschool (max)-preschool (min):		
1	83.81%	70.48%
2	11.82%	21.69%
3	4.38%	7.83%
Share of children who spent the most years in preschool in a family by birth order:		
1	24.89%	28.40%
2	57.21%	59.26%
3	15.72%	9.88%
4	2.18%	2.47%
Number of observations	1195	365

The sibling sample consists of children from families in which at least two children provide information on preschool attendance during the three years prior to school start and on secondary school track attendance. *Share of children by within-family difference in preschool attendance (in years)* indicates the share of children who live in families with the respective difference in years of preschool attendance (only children from families in which preschool attendance differs across siblings are considered). *Share of children who spent the most years in preschool in a family by birth order* indicates the share of children by birth order who have the longest exposure to preschool in their families (only children from families in which preschool attendance differs across siblings are considered).

Table 5
Sibling Sample: Preschool Effects with and without Family Fixed Effects

	Full sibling sample				Sibling sample of children with migration background		
	Full sample	Full sample	Parents < upper secondary	Parents < middle school	Full sample	Parents < upper secondary	Parents < middle school
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A) Sibling sample **without** family fixed effects:							
Years of preschool attendance	0.071***	0.035*	0.060***	0.073***	0.063**	0.084***	0.086***
	(0.019)	(0.018)	(0.020)	(0.025)	(0.027)	(0.029)	(0.024)
B) Sibling sample **with** family fixed effects:							
Years of preschool attendance	–0.033	–0.023	–0.019	–0.020	0.027	0.036	0.027
	(0.023)	(0.022)	(0.024)	(0.024)	(0.028)	(0.030)	(0.024)
Family fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,195	1,195	838	427	365	319	264
C) Sibling sample **with** family fixed effects, only families with variation in preschool attendance across siblings:							
Years of preschool attendance	–0.033	–0.020	–0.010	–0.023	0.019	0.018	0.013
	(0.023)	(0.022)	(0.026)	(0.026)	(0.030)	(0.031)	(0.026)
Control variables	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	457	457	330	174	166	148	116

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns show results from linear probability models. *Control variables* include the following covariates: highest educational degree of parents (only in specifications without family fixed effects), sex, employment status of the mother (in the year of entrance into preschool), household income (in the year of entrance into preschool), birth order, age at school entrance, age at track entrance, dummies for year of track entrance, and dummies for federal state of track entrance. *Parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *Parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 6
Sibling Sample: Preschool Effects with and without Family Fixed Effects—Effect on a Higher Secondary School Attendance than Basic School (Hauptschule)

	Full sibling sample				Sibling sample of children with migration background		
	Full sample	Full sample	Parents < upper secondary	Parents < middle school	Full sample	Parents < upper secondary	Parents < middle school
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A) Sibling sample **without** family fixed effects:							
Years of preschool attendance	0.058***	0.045**	0.067***	0.062*	0.042	0.046	0.049
	(0.022)	(0.020)	(0.025)	(0.036)	(0.027)	(0.033)	(0.034)
B) Sibling sample **with** family fixed effects:							
Years of preschool attendance	–0.010	0.004	0.028	–0.000	0.012	0.012	0.003
	(0.023)	(0.028)	(0.034)	(0.033)	(0.036)	(0.040)	(0.042)
Family fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,195	1,195	838	427	365	319	264

Dependent variable: Secondary school track higher than basic school (Hauptschule) attendance (1=yes, 0=no). All columns show results from linear probability models. *Control variables* include the following covariates: highest educational degree of parents (only in specifications without family fixed effects), sex, employment status of the mother (in the year of entrance into preschool), household income (in the year of entrance into preschool), birth order, age at school entrance, age at track entrance, dummies for year of track entrance, and dummies for federal state of track entrance. *Parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *Parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 7
Effect Heterogeneity: More than 1 Year Preschool Difference across Siblings

	Full sibling sample			Sibling sample of children with migration background		
	Full sample	Parents < upper secondary	Parents < middle school	Full sample	Parents < upper secondary	Parents < middle school
	(1)	(2)	(3)	(4)	(5)	(6)
A) Sibling sample **without** family fixed effects:						
Years of preschool attendance	0.062***	0.090***	0.099***	0.057*	0.083**	0.075***
	(0.021)	(0.023)	(0.029)	(0.030)	(0.032)	(0.026)
B) Sibling sample **with** family fixed effects:						
Years of preschool attendance	0.032	0.032	0.017	0.025	0.034	0.013
	(0.023)	(0.024)	(0.022)	(0.034)	(0.038)	(0.031)
Family fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	806	568	287	243	210	178

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns show results from linear probability models. We dropped all families with two children and a one-year difference in preschool attendance across siblings. In families with three or four children, we dropped children with two years of attendance (if within-family preschool attendance differs between one and three years or zero and three years) and with one year of attendance (if within-family preschool attendance differs between zero and two years). *Control variables* include the following covariates: highest educational degree of parents (only in specifications without family fixed effects), sex, employment status of the mother (in the year of entrance into preschool), household income (in the year of entrance into preschool), birth order, age at school entrance, age at track entrance, dummies for year of secondary school track entrance, and dummies for federal state of secondary school track entrance. *Parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *Parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 8
Effect Heterogeneity: Preschool Effects with and without Family Fixed Effects - Interaction with Full-Day Attendance

	Full sibling sample				Sibling sample of children with migration background		
	Full sample	Full sample	Parents < upper secondary	Parents < middle school	Full sample	Parents < upper secondary	Parents < middle school
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A) Sibling sample **without** family fixed effects:							
Years of preschool attendance	0.063***	0.021	0.054**	0.063**	0.048	0.076**	0.058**
	(0.020)	(0.019)	(0.021)	(0.028)	(0.031)	(0.032)	(0.028)
Interaction with full-day attendance	0.085***	0.042	0.023	0.018	0.046	0.019	0.055
	(0.032)	(0.031)	(0.030)	(0.033)	(0.042)	(0.042)	(0.039)
B) Sibling sample **with** family fixed effects:							
Years of preschool attendance	–0.042	–0.028	–0.006	–0.012	0.044	0.052	0.032
	(0.026)	(0.025)	(0.028)	(0.028)	(0.035)	(0.037)	(0.036)
Interaction with full-day attendance	0.024	0.002	–0.016	–0.005	–0.008	–0.018	0.018
	(0.030)	(0.033)	(0.038)	(0.033)	(0.043)	(0.046)	(0.039)
Family fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,195	1,195	838	427	365	319	264

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns show results from linear probability models. *Interaction with full-day attendance* is the interaction term between *years of preschool attendance* and *years of full-day preschool attendance*. *Control variables* include the following covariates: highest educational degree of parents (only in specifications without family fixed effects), sex, employment status of the mother (in the year of entrance into preschool), household income (in the year of entrance into preschool), birth order, age at school entrance, age at track entrance, dummies for year of track entrance, and dummies for federal state of track entrance. *Parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *Parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 9
Effect Heterogeneity: Children with Migration Background and Mothers' Language

	Sibling sample of children with migration background			
	Full sample		Parents <upper secondary	Parents <middle school
	(1)	(2)	(3)	(4)
Years of preschool attendance	–0.071	–0.036	–0.031	–0.047
	(0.044)	(0.060)	(0.070)	(0.064)
Years of preschool attendance × Mother speaks no German	0.103**	0.087	0.081	0.079
	(0.051)	(0.067)	(0.075)	(0.065)
Employment status of the mother		Yes	Yes	Yes
Household income		Yes	Yes	Yes
Sex		Yes	Yes	Yes
Birth order		Yes	Yes	Yes
Age at school entrance		Yes	Yes	Yes
Age at track entrance		Yes	Yes	Yes
Dummies for year of secondary school track entrance		Yes	Yes	Yes
Dummies for federal state of secondary school track entrance		Yes	Yes	Yes
Observations	344	344	303	252

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns report results from linear probability models with family fixed effects. *Mother speaks no German* is a dummy variable that equals 1 if the mothers speaks mostly or partly her language of origin and 0 if the mother speaks mostly German. The few cases missing this variable lead to the slightly smaller sample size compared to the full sample of children with migration background. *Parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *Parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 10
Siblings' Differences in Innate Ability and Social Skills

	Whole sample	Sibling sample
A) Innate ability (Scale 0-20):		
Mean innate ability	9.92	10.26
SD innate ability	3.51	3.16
Average SD of innate ability within families		2.06
Average SD of innate ability within families:		
No variation in preschool attendance across siblings		2.52
Variation in preschool attendance across siblings		1.32
Number of observations	414	135
B) Social skills (Scale 0-5):		
Mean social skills	4.03	4.01
SD social skills	1.12	1.11
Average SD of social skills within families between siblings		0.70
Average SD of social skills within families between siblings:		
No variation in preschool attendance between siblings		0.70
Variation in preschool attendance between siblings		0.70
Number of observations	1132	306

Panel A: The whole sample consists of all children who provide information on preschool attendance (in the three years before school start), secondary school track attendance, and innate ability. The sibling sample provides the same information for at least two children of one family. Cognitive ability is assessed at age 17 when children for the first time themselves answer the GSOEP household questionnaire. The test aims at measuring so-called *fluid cognitive competencies* that capture general, less education-specific, clearly genetically or biologically affected ability. *Fluid cognitive abilities* are, compared to crystalline cognitive abilities, less culture- and lifecycle-specific and more stable across educational and professional pathways (Weinert et al. 2007). The assessment consists of three parts with 20 questions each (verbal, numerical, and figural competencies). We use only the subscale of the figural test since these competencies are most consistent with the concept of *fluid cognitive competencies* and measure innate ability. Adolescents were assessed between 2006 and 2009.

Panel B: Data are from GSOEP's mother-child data set on three year old children (assessed for children born between 2002 and 2006). The whole sample consists of all three year old children with information on social skills, age (in months), and preschool attendance. The sibling sample provides the same information for at least two children of one family. Mothers assess their children's social skills on the basis of the following five statements: (I) child calls familiar people by name, (II) child plays with other children, (III) child participates in role-playing games, (IV) child has preferences for certain friends, and (V) child calls own feelings by name. Mothers can answer *yes*, *to some extent*, or *no*. The number of *yes* answers (1-5) is our social skills outcome.

Table 11
Robustness Check: Controlling for Innate Ability

	Sibling sample			
	(1)	(2)	(3)	(4)
Years of preschool attendance	0.065	0.080	–0.031	–0.056
	(0.086)	(0.085)	(0.115)	(0.115)
Innate ability		0.031**		–0.022
		(0.015)		(0.020)
Highest educational degree of parents	Yes	Yes	No	No
Employment status of mother	Yes	Yes	Yes	Yes
Household income	Yes	Yes	Yes	Yes
Sex	Yes	Yes	Yes	Yes
Birth order	Yes	Yes	Yes	Yes
Age at school entrance	Yes	Yes	Yes	Yes
Age at track entrance	Yes	Yes	Yes	Yes
Dummies for year of secondary school track entrance	Yes	Yes	Yes	Yes
Dummies for federal state of secondary school track entrance	Yes	Yes	Yes	Yes
Family fixed effects	No	No	Yes	Yes
Observations	135	135	135	135

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns report results from linear probability models. *Innate ability* is the child's performance on a test of figural skills taken at age 17. We use the number of correct answers on this test (scale from 0-20). We include only children with information on preschool attendance in the three years before school start, on secondary school track attendance, and on innate ability from the figural skills test. Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 12
Robustness Check: Correlation between Social Skills and Preschool Attendance

	Whole sample		Sibling sample	
	(1)	(2)	(3)	(4)
Social Skills	0.083***	0.076***	0.072***	0.040
	(0.013)	(0.013)	(0.025)	(0.025)
Age of the child (in months)		0.013***	0.007	0.009
		(0.004)	(0.008)	(0.010)
Family fixed effects	No	No	No	Yes
Observations	1,132	1,132	306	306

Dependent variable: Preschool attendance (1=yes, 0=no). All columns report results from linear probability models. *Social skills* is a measure for children's social skills at age three. Mothers assess their children's social skills by responding to the following five statements: (I) child calls familiar people by name, (II) child plays with other children, (III) child participates in role-playing games, (IV) child has preferences for certain friends, and (V) child calls own feelings by name. Mothers can answer *yes*, *to some extent*, or *no*. The number of *yes* answers (1-5) is our measure for social skills. Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 13
Cognitive Skills as Dependent Variable

	Sibling sample			
	(1)	(2)	(3)	(4)
Years of preschool attendance	–0.067	–0.142	–1.200	–0.323
	(1.180)	(1.450)	(1.548)	(1.500)
Highest educational degree of parents		Yes		
Employment status of mother		Yes		Yes
Household income		Yes		Yes
Sex		Yes		Yes
Birth order		Yes		Yes
Age at school entrance		Yes		Yes
Age at track entrance		Yes		Yes
Dummies for year of secondary school track entrance		Yes		Yes
Dummies for federal state of secondary school track entrance		Yes		Yes
Family fixed effects	No	No	Yes	Yes
Observations	137	137	137	137

Dependent variable: Number of correct answers on cognitive skills tests assessing numerical and verbal skills of 17-year-old adolescents (on a scale from 0-40). We include only children with information on preschool attendance in the three years before school start and on cognitive skills from both the numerical and verbal skills test. Robust standard errors (in parentheses) clustered at the family level. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A1
Sibling Sample: Preschool Effects with and without Family Fixed Effects
(Omitting Children attending Gesamtschule)

	Full sibling sample				Sibling sample of children with migration background		
	Full sample	Full sample	Parents < upper secondary	Parents < middle school	Full sample	Parents < upper secondary	Parents < middle school
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A) Sibling sample **without** family fixed effects:							
Years of preschool attendance	0.073***	0.053***	0.065***	0.072**	0.068**	0.088***	0.081***
	(0.021)	(0.019)	(0.022)	(0.030)	(0.031)	(0.033)	(0.030)
B) Sibling sample **with** family fixed effects:							
Years of preschool attendance	–0.025	–0.026	–0.019	–0.023	0.011	0.025	0.026
	(0.025)	(0.025)	(0.027)	(0.027)	(0.028)	(0.034)	(0.027)
Family fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,070	1,070	756	381	327	283	264

Dependent variable: High school (Gymnasium) attendance (1=yes, 0=no). All columns show results from linear probability models. *Control variables* include the following covariates: highest educational degree of parents (only in specifications without family fixed effects), sex, employment status of the mother (in the year of entrance into preschool), household income (in the year of entrance into preschool), birth order, age at school entrance, age at track entrance, dummies for year of secondary school track entrance, and dummies for federal state of secondary school track entrance. *Parents<upper secondary* means that parents have no more than a degree from middle school (Realschule). *Parents<middle school* means that parents have no more than a degree from basic school (Hauptschule). Robust standard errors (in parentheses) clustered at the family level. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Ifo Working Papers

No. 105 Sinn, H.-W. und T. Wollmershäuser, Target-Kredite, Leistungsbilanzsalden und Kapitalverkehr: Der Rettungsschirm der EZB, Juni 2011.

No. 104 Czernich, N., Broadband Internet and Political Participation: Evidence for Germany, June 2011.

No. 103 Aichele, R. and G. Felbermayr, Kyoto and the Carbon Footprint of Nations, June 2011.

No. 102 Aichele, R. and G. Felbermayr, What a Difference Kyoto Made: Evidence from Instrumental Variables Estimation, June 2011.

No. 101 Arent, S. and W. Nagl, Unemployment Benefit and Wages: The Impact of the Labor Market Reform in Germany on (Reservation) Wages, June 2011.

No. 100 Arent, S. and W. Nagl, The Price of Security: On the Causality and Impact of Lay-off Risks on Wages, May 2011.

No. 99 Rave, T. and F. Goetzke, Climate-friendly Technologies in the Mobile Air-conditioning Sector: A Patent Citation Analysis, April 2011.

No. 98 Jeßberger, C., Multilateral Environmental Agreements up to 2050: Are They Sustainable Enough?, February 2011.

No. 97 Rave, T., F. Goetzke and M. Larch, The Determinants of Environmental Innovations and Patenting: Germany Reconsidered, February 2011.

No. 96 Seiler, C. and K. Wohlrabe, Ranking Economists and Economic Institutions Using RePEc: Some Remarks, January 2011.

No. 95 Itkonen, J.V.A., Internal Validity of Estimating the Carbon Kuznets Curve by Controlling for Energy Use, December 2010.

No. 94 Jeßberger, C., M. Sindram and M. Zimmer, Global Warming Induced Water-Cycle Changes and Industrial Production – A Scenario Analysis for the Upper Danube River Basin, November 2010.

No. 93 Seiler, C., Dynamic Modelling of Nonresponse in Business Surveys, November 2010.

No. 92 Hener, T., Do Couples Bargain over Fertility? Evidence Based on Child Preference Data, September 2010.

No. 91 Schlotter, M. und L. Wößmann, Frühkindliche Bildung und spätere kognitive und nicht-kognitive Fähigkeiten: Deutsche und internationale Evidenz, August 2010.

No. 90 Geis, W., High Unemployment in Germany: Why do Foreigners Suffer Most?, August 2010.

No. 89 Strobel, T., The Economic Impact of Capital-Skill Complementarities in German and US Industries – Productivity Growth and the New Economy, July 2010.

No. 88 Falck, O., M. Fritsch and S. Heblich, The Phantom of the Opera: Cultural Amenities, Human Capital, and Regional Economic Growth, June 2010.

No. 87 Strobel, T., Institutions and Innovations as Sources of Productivity Growth – Cross-Country Evidence, April 2010.

No. 86 Strobel, T., Unraveling the Origins of EU Countries Productivity Growth – Evidence on R&D and Competition from Cross-Country Industry Analysis, April 2010.

No. 85 Podlich, N., D. Illyasov, E. Tsoy and S. Shaikh, The Methodology of Stress Tests for the Kazakh Banking System, April 2010.

No. 84 Jaeger, U., Working or stay-at-home mum? The Influence of family benefits and religiosity, March 2010.

No. 83 Montén, A. and C. Thater, Determinants of Efficiency in Child Care Provision, March 2010.

No. 82 Ebertz, A., The Determinants of Joint Residential and Job Location Choice: A Mixed Logit Approach, December 2009.

No. 81 Gronwald, M., J. Mayr and S. Orazbayev, Estimating the Effects of Oil Price Shocks on the Kazakh Economy, October 2009.

No. 80 Geis, W., Does Educational Choice Erode the Immigration Surplus?, October 2009.