Efficiency and Equity of Education Tracking
A Quantitative Analysis*

Suzanne Bellue† Lukas Mahler ‡

October 9, 2023

[Click here for the latest version]

Abstract

We study the long-run aggregate, distributional, and intergenerational effects of school tracking—the allocation of students to different types of schools—by incorporating school track decisions into a general-equilibrium heterogeneous-agent overlapping-generations model. The key innovation in our model is the school skill production technology with school tracking. School tracks endogenously differ in their pace of instruction and the students' average skills. We show analytically that this technology can rationalize reduced-form evidence on the effects of school tracking on the distribution of end-of-school skills. We then calibrate the model using representative data from Germany, a country with a very early school tracking policy by international standards. Our calibrated model shows that an education reform that postpones the tracking age from ten to fourteen generates improvements in intergenerational mobility but comes at the cost of modest losses in aggregate human capital and economic output, reducing aggregate welfare. This efficiency-mobility trade-off is rooted in the effects of longer comprehensive schooling on learning and depends crucially on the presence of general equilibrium effects in the labor market. Finally, counterfactual analyses suggest that policies that reduce the parental influence in the school track choice can increase both social mobility and aggregate economic output, improving aggregate welfare.

Keywords: Intergenerational Mobility, Education Tracking, Inequality, Efficiency

JEL Codes: E24, I24, J24

*We are very grateful to Antonio Ciccone, Michèle Tertilt, and Minchul Yum for their continued and invaluable support and guidance in this project. In addition, we thank Hans Peter Grüner, Andreas Gulyas, Yannick Reichlin, and numerous seminar and conference for helpful discussions and suggestions. Suzanne Bellue gratefully acknowledges financial support from the German Academic Exchange Service (DAAD) and the German Research Foundation (through the CRC-TR-224 project A03). Lukas Mahler gratefully acknowledges the financial support from the German Science Foundation (through the CRC TR 224 Project A04) and the SFB 884 Political Economy of Reforms.

†CREST-ENSAE (suzanne.bellue@ensae.fr)

‡University of Mannheim (lukas.mahler@gess.uni-mannheim.de)

1 Introduction

School tracking—the allocation of school children into different types of schools at some point during their school career—is a common feature of education policy across OECD countries. For example, in 2012, 20 out of 34 countries in the OECD had an education system with at least two distinct school programs available to 15-year-old students.[1] The argument behind tracking is typically one of efficiency: grouping children by ability and aspirations creates more homogeneous classrooms and allows for tailored instruction levels and curricula, which improves educational outcomes (Duflo et al., 2011). On the other hand, as the track decision is often related to the socioeconomic background of children, tracking may impair socioeconomic mobility across generations and increase inequality in education and income (Carlana et al., 2022; Meghir and Palme, 2005; Pekkarinen et al., 2009; Hanushek and Wössmann, 2006). This concern is particularly strong for countries that track at a relatively young age (Dustmann, 2004). For this reason, school tracking, and, in particular, its timing, is a recurrent issue in the public and academic debate about education reforms in countries with an early tracking regime, such as Germany (OECD, 2020).[2]

This paper contributes to the debate by quantitatively assessing the long-run aggregate, distributional, and inter-generational effects of school tracking policies. Any such assessment needs to consider the effects of tracking on educational outcomes in school and in college, the effect of the supply of different skills on labor market outcomes, and the intergenerational effects of parental skills differences. Quantitative macroeconomic models of overlapping generations have proven useful in analyzing these effects, and the interplay between them, but have so far not incorporated how skill accumulation is affected by school tracking policies (Lee and Seshadri, 2019; Daruich, 2022; Restuccia and Urrutia, 2004; Yum, 2023). We aim to fill this gap by providing a macroeconomic model that features tracking in secondary school, allowing us to quantify the role that tracking plays for aggregate and distributional

[1]An overview of school tracking policies in OECD countries is given in Chapter 2 in OECD (2013). We differentiate school tracking, which refers to allocating students into physically distinct types of schools that differ in the curriculum taught, intensity, and length, from ability grouping within a school, where the curriculum and educational goals remain the same. School tracking is also common among non-OECD countries. Based on 2012 PISA data, only two out of 26 non-OECD countries with available information featured an education system with one comprehensive school program available to 15-year-old students.

[2]There is substantial variation in the timing of tracking across OECD countries (see Figure IV.2.4 in OECD, 2013, based on 2012 PISA data). Germany and Austria are among the countries with the earliest track selection, at age 10. Turkey, Hungary, and the Czech Republic track at age 11; Belgium, the Netherlands, and Switzerland at age 12; Luxembourg at age 13; Italy, South Korea, and Slovenia at age 14; France, Greece, Israel, Mexico, Portugal and many other countries at age 15. Other countries, like the US, UK, and Australia, do not track during secondary school.

socioeconomic outcomes, within and across generations.

The model is built around a parsimonious theory of how school tracking affects the accumulation of skills in school. Our skill accumulation technology implies that each child has an optimal pace of instruction, which is increasing in children's skills. The pace of instruction is allowed to differ across school tracks but not within tracks. Policymakers choose the pace of instruction across tracks to maximize aggregate skills at the end of secondary school. We also allow for direct peer effects—children learn more if their school peers have higher skills. Because of the (endogenous) stratification in skills implied by school tracking, this is a further channel through which tracking affects skill accumulation in school. Under linear direct peer effects and absent any shocks to child skills during their time in school, the skill formation technology implies that the optimal tracking policy should perfectly stratify children when they start school. However, in the more realistic scenario where children's skills develop at different—and hard to predict—tempos as they grow older, early tracking may lead to lower aggregate end-of-school skills because of a mismatch between children's skills and the pace of instruction. Also, early tracking can increase inequality in educational outcomes. Another interesting implication of tracking is that children who lose in terms of skills are often concentrated in the track with the lower instruction pace. Thus, our child skill formation technology rationalizes some of the most robust empirical findings regarding school tracking in the literature and encompasses the main arguments about school tracking frequently made in the public discourse.[3]

We embed our theory of skill accumulation in school into a general equilibrium life-cycle incomplete-markets framework of overlapping generations, in which parents care about their offspring in the tradition of Becker and Tomes (1986). Some aspects of the model are tailored to fit the German Education System. Children are tracked into two school tracks at the age of ten (typically at the end of four years of primary school) based on their parents' decisions. While predominantly based on child skills, the actual track choice may also reflect parental preferences over school tracks. While only one of the two tracks leads directly to college, there is a second-chance opportunity for children in the other track.

[3]Empirical estimates of the effects of (early) tracking on the average learning outcomes of school children are often ambiguous (Hanushek and Wössmann, 2006). Evidence for the effects of tracking on inequality is more consistent, finding that tracking raises educational inequality and tends to predominantly disadvantage children from lower socioeconomic backgrounds (see, for instance, Meghir and Palme (2005), Aakvik et al. (2010), and Pekkala Kerr et al. (2013) for evidence from Scandinavian countries and Matthewes (2021); Piopiunik (2014), for the case of Germany). While opponents of early tracking argue in favor of postponing the tracking age as a means to increase equality of opportunity in access to education for disadvantaged children (for example Woessmann (2020) in Germany), proponents argue that in a comprehensive school, faster learning children are thwarted, while slower-learning children overstrained, resulting in learning losses (see, for example Esser and Seuring (2020) in Germany).

Going to college allows access to the college-skilled labor market and affects human capital growth rates, but incurs psychic costs, which are a function of child skills, as well as time costs relative to non-college education. End-of-school child skills translate into adult human capital, which evolves stochastically over the working life and determines, together with the college education decision, labor earnings. The distribution of human capital across college and non-college workers affects equilibrium wages, which are anticipated when parents choose the school track for their children. Households can save into a non-state-contingent asset subject to life-cycle borrowing constraints. Finally, parents can make a non-negative inter-vivos transfer when children become independent.

We solve for the steady-state equilibrium of the model numerically and calibrate the parameters in two steps. First, we estimate the child skill formation technology parameters directly from German data on school children (Blossfeld et al., 2019) using a latent variable framework as in Cunha et al. (2010); Agostinelli et al. (2023). In particular, we use information on achievement test scores to measure child skills at different stages of their school careers. We then calibrate the remaining parameters to match a set of salient moments from representative German survey data. The model matches the data well, both in terms of aggregate moments and in terms of the distribution of child skills across school tracks and parental backgrounds, as well as the transitions through the education system. To test the model's validity, we investigate non-targeted moments, such as the determinants of the school track choice. The model reproduces well the relationship between skills and school track choice by parental background. In addition, we also check our calibrated model against Dustmann et al. (2017)'s empirical observation that for children at the margin between the two tracks in Germany, school tracking is inconsequential for earnings later in life. We compute the effects of the initial school track on later-in-life economic outcomes for a set of children who are, in equilibrium, just at the margin between the two school tracks. Simulated data from our model confirms that very similar children at the time of the track decision who end up in different school tracks experience similar lifetime economic outcomes.

A variance decomposition exercise shows that skill formation during the school tracking years plays an essential role for lifetime inequality across the population. In particular, around a third of the variation in lifetime economic outcomes is accounted for at age ten, just after the school track choice. This share rises to around two-thirds at age eighteen after the college choice, a number consistent with the literature. This suggests that the evolution of skills during the tracking years in secondary school is crucial for determining lifetime inequality, underscoring the importance of understanding skill formation during these years.

Our main findings regard the long-run and welfare effects of counterfactual policies that change the timing of school tracking. We find that a policy reform that postpones the school tracking age by four years to age 14—the average tracking age in OECD countries—entails an efficiency-mobility trade-off. On the one hand, postponing the tracking age improves social mobility as it decreases the intergenerational elasticity of income by around 2.2%. These mobility gains arise primarily because the longer time spent in comprehensive school decreases the overall heterogeneity in end-of-school skills, as those who would have gone to an academic track school lose while those who would have gone to a vocational track school gain. As the track choice is correlated with parental education, children from different parental backgrounds also become more similar, and the later track choice depends less on parental background. As a result, end-of-school skills, college, and labor market outcomes become less dependent on the initial skills children inherit from their parents, improving mobility. At the same time, smaller heterogeneity in end-of-school skills also translates into smaller differences in human capital across college and non-college workers. Consequently, the college wage premium in efficiency units falls, and inequality, as measured by the Gini coefficient of earnings, drops by 0.4%.

On the other hand, our results indicate that postponing tracking comes at the cost of a 0.1% drop in GDP. This is because prolonged learning in a comprehensive school track foregoes efficiency gains from tailored instruction levels in an early tracking system. Quantitatively, these learning losses cannot be recuperated even though the later tracking decision occurs after more uncertainty about skills has been resolved and can therefore be more accurate. Lower levels of child skills then translate into lower levels of human capital in the economy, causing lower aggregate output. In this context, we highlight the importance of considering general equilibrium effects on the labor market that influence school track and college decisions. In partial equilibrium, when wages are fixed, aggregate output would even increase as the share of college-educated workers rises considerably. Driven by the efficiency losses, our welfare analysis suggests that a later tracking system would decrease average welfare by around 0.05% in consumption equivalent units. Abolishing tracking in favor of comprehensive schooling altogether further exacerbates the efficiency-mobility trade-off.[4]

Finally, we evaluate the effects of reducing the direct influence of a child's socioeconomic background on the school track choice without modifying the timing of school tracking. Our

[4] A similar trade-off has been highlighted in the literature about the effects of economic segregation on growth and inequality (see Benabou, 1996) and more recently by Arenas and Hindriks (2021) where efficiency gains from unequal school opportunities arise because of positive assortative matching between parents who invest more in their children and better schools. In contrast, in our case, efficiency gains arise from matching similar-ability peers to tailored instruction levels for longer.

data indicate that when parents go against the school track recommendation of their children's primary school teachers, it is generally in favor of the school track parents attended in the past. We rationalize this as coming from parental preferences in favor of their own school tracks.[5] These parental preferences turn out to play an important role in children's school track choice in our calibrated model. This is consistent with previous findings in the literature (e.g. Dustmann, 2004), and may give rise to inefficiencies in the allocation of children across tracks.[6] An important question is whether the consequences of such "misallocation" effects are visible not only in child skill outcomes but also in the aggregate and distributional outcomes in the economy. We show that reducing the direct influence of parental background on the school track by reducing parental preferences for their own school track to zero leads to improvements in both social mobility and economic output in the range of 0.9% and 0.04%, respectively, improving aggregate welfare by 0.04%. These improvements result from the fact that the initial school track choice becomes less driven by parental economic background and more on skills, which improves the teaching efficiency in each track and thereby raises the average skill level. These results highlight that measures, such as mentoring programs, which provide information and support to children from lower socioeconomic families and have been effective in alleviating the negative influence of family background on school track decisions (Raposa et al., 2019), can not only improve the outcomes of these individual children but also lead to aggregate efficiency gains in the economy.

Related Literature

This paper links several strands of the literature: the quantitative macroeconomic literature on inequality and mobility, the literature on children's skill formation during school years, and the school tracking literature.

Much research in the the quantitative macroeconomic literature on inequality and intergenerational mobility shares our focus on the role of skill formation, education, and education policies (e.g. Becker and Tomes, 1979, 1986; Restuccia and Urrutia, 2004; Lee and Seshadri,

[5]Those preferences may refer to actual track tastes, as parents might want their children to follow in their footsteps (Doepke and Zilibotti, 2017), or to some form of biased information about how each school track influences later outcomes or about the costs associated with completing it. We cannot tell the underlying reasons apart.

[6]For example, a college-educated parent may push her child into an academic-track school even though her child's skills optimally suggest a vocational-track school. This harms her child's learning outcomes and affects average learning in that track as the instruction pace endogenously adjusts to the composition of skills in that track. We calibrate the extent of these asymmetric preferences in our model to replicate the share of deviations of the chosen school track from what had been recommended by the primary school teachers in our data.

2019; Abbott et al., 2019). However, while many papers in this literature focus on early childhood education and the role of parental investments (e.g. Daruich, 2022; Yum, 2023; Caucutt and Lochner, 2020; Lee and Seshadri, 2019), or study the role of college-education policies (e.g. Krueger and Ludwig, 2016; Abbott et al., 2019; Capelle, 2022), few papers explicitly include the secondary schooling stage into their analysis. An exception is Fujimoto et al. (2023), who study the importance of free secondary schooling for misallocation driven by borrowing constraints in Ghana. However, in their context, secondary schooling is the highest education level. In addition, recent research has highlighted the heterogeneous impact of school closures in the wake of the Covid pandemic on children at different stages of their schooling career (Jang and Yum, 2022; Fuchs-Schündeln et al., 2022) and across public and private secondary schools (Fuchs-Schündeln et al., 2023). Our contribution is to study a widespread feature of education policy during secondary school – tracking. Despite the fact that tracking occurs in many countries and that (early) tracking policies are often made responsible for persistent inequality and social immobility in the public debate, an extensive analysis of broad reforms to tracking policies in the macroeconomic literature is, to the best of our knowledge, missing.[7]

Our theory of skill formation during schooling years builds on the insights of the literature on child skill formation, which studies how children's skills evolve as a function of endowments, parental and environmental inputs, and schooling and teaching inputs (see, for instance, Cunha and Heckman, 2007; Cunha et al., 2010; Agostinelli et al., 2023, 2019; Duflo et al., 2011; Aucejo et al., 2022; Bonesrønning et al., 2022).[8] To incorporate how tracking affects learning in secondary school, we consider two forms of peer effects. First, similar to Agostinelli (2018), we incorporate direct peer effects, which capture the idea that children are affected by different-quality peer groups in a school track. Second, following Duflo et al. (2011)'s evidence in Kenyan primary schools and Aucejo et al. (2022)'s findings of comple-

[7]There is an extensive literature in education economics, which theoretically analyzes tracking policies (see Epple et al. (2002) and Betts (2011) for a general theoretical foundation of tracking), to which we relate. This literature tends to conclude that the effects of tracking on the level and distribution of educational outcomes are often theoretically ambiguous and depend on the shape of peer effects, resources between tracks, or the uncertainty surrounding child abilities. We make a similar point in Section 3. Brunello et al. (2007) offer an analysis of the optimal timing of tracking focusing on the role that an increasing demand towards more general skills plays, while Brunello et al. (2012) estimate the efficiency losses of deviating from the optimal tracking age across Europe, finding losses in the range of half a percent of GDP, on average. Our contribution to this literature is that we provide a richer framework that incorporates important macroeconomic effects of tracking on higher education and labor market outcomes and allows us to draw conclusions on the effects of tracking on mobility across generations.

[8]Sacerdote (2011) provides an overview of empirical approaches to measuring peer effects in education. A summary of theoretical models of peer interactions and their implications for tracking policies can be found in Epple and Romano (2011).

mentarities between classroom composition and teaching practice in the U.S., we consider how the instruction levels across tracks adjust endogenously to the skill composition in that track. For that reason, school tracking is conceptually different from schools of different qualities (often related to neighborhood effects), which would mechanically disadvantage children in the lower-quality tracks. For example, Arenas and Hindriks (2021) provide a model of unequal school opportunity, defined as unequal school quality and access probability to the best schools, and quantify its effect on intergenerational persistence, highlighting the role of positive assortative matching between parents who invest more into their children and high-quality schools. In contrast, schools tracks differ endogenously in their instruction pace. Choosing a school track is thus less about choosing a "good" versus a "bad" school, but more about choosing a school that fits the learning needs of a child.[9]

We embed the skill formation technology with tracking during secondary school into a standard life-cycle model with intergenerational linkages, such that the intital conditions of a new generation are endogenous, following the work in Daruich (2022); Lee and Seshadri (2019); Yum (2023). Moreover, we share with these papers the importance of considering GE effects when studying policy reforms that affect the skill composition in the economy. While the quantitative macroeconomic literature focuses almost exclusively on the US, where tracking across schools is uncommon (Fujimoto et al., 2023, is a notable exception as they focus on a developing country, Ghana), we focus on a country with a very early tracking system—Germany. At the same time, our school track model is general enough to be used in other countries that track between schools, such as many European countries, but also Asian countries like Korea and Singapore, or South American countries like Uruguay (OECD, 2013), and could even be adapted to be informative for countries where tracking occurs mostly within schools, across classrooms, such as in the US.

Lastly, this paper connects to an extensive empirical literature that estimates the effects of school tracking, and in particular its timing, on educational and later-in-life outcomes of students (see Betts, 2011, for an excellent overview). This literature typically either exploits temporal within-country variation in tracking practices (Meghir and Palme (2005), for Sweden; Aakvik et al. (2010), for Norway; Bauer and Riphahn (2006), for Switzerland; Malamud and Pop-Eleches (2011), for Romania; Pekkala Kerr et al. (2013), for Finland; and Matthewes (2021); Piopiunik (2014) for Germany) or between-country variation with a difference-in-differences strategy (Hanushek and Wössmann, 2006; Ruhose and Schwerdt,

[9]Similarly, tracking is also different from having private schools or schools with different costs, where selection into schools is likely to depend on parental wealth directly. The degree to which private schools are similar to a tracking system will then depend on the correlation of child skills and parental wealth.

2016). Most studies suggest that earlier tracking raises inequality in educational outcomes and increases the effect of parental education on student achievement. A notable exception is Dustmann et al. (2017), who use an individual-level instrumental variables strategy (the date of birth) and find no effect of track choice on educational attainment or earnings for students at the margin between two tracks. This result suggests that school tracking in Germany is inconsequential in the longer run for children whose skills put them in between tracks at the time the decision was made, a result that our model replicates. We add to this literature a quantitative model-based assessment of the long-term aggregate, distributional, and welfare effects of broad reforms to the school tracking age, which is difficult to establish empirically.

The remainder of the paper is organized as follows. Section 2 presents our model of overlapping generations and tracking during secondary school and introduces the child skill formation technology. In Section 3, we build intuition about the model mechanisms underlying school tracking by deriving theoretical implications of that technology. Section 4 explains how we parameterize and calibrate the model. It also offers some validation exercises. In Section 5, we use the calibrated model to perform a series of counterfactual experiments to quantify the effects of different school tracking policy regimes. Finally, Section 6 concludes.

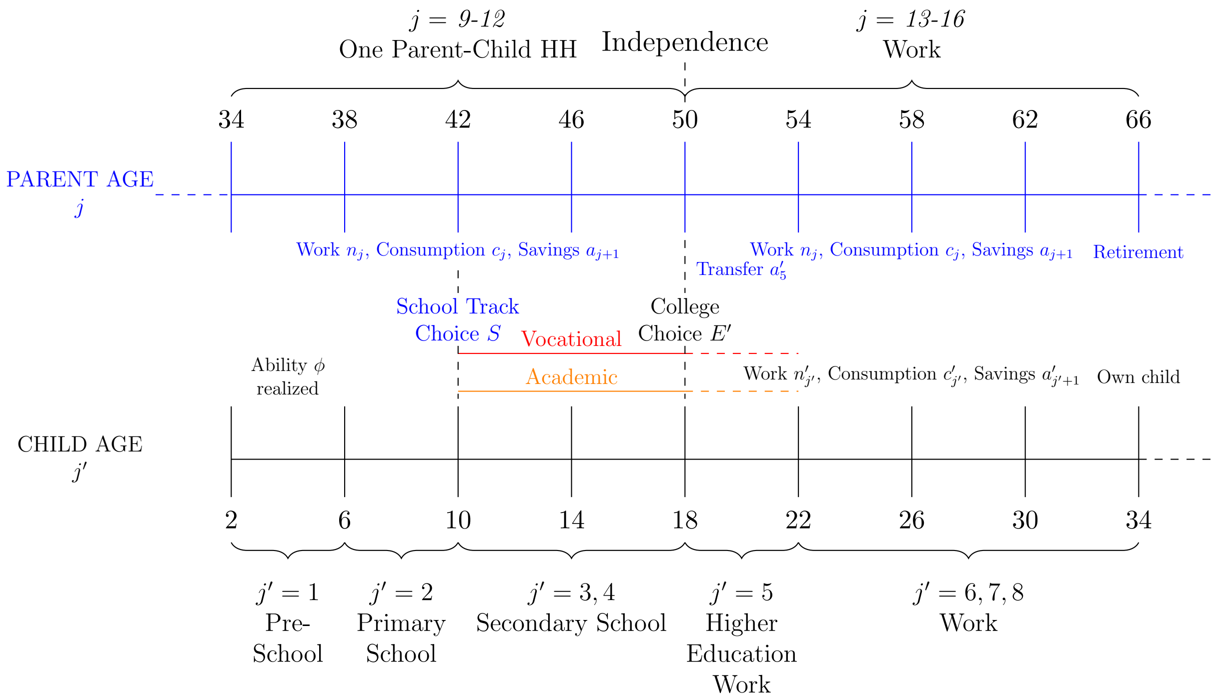
2 The Model

Time is discrete and infinite, and one model period, $j \in \{1, ..., 20\}$, corresponds to the four years between ages $[4j - 2, 4j + 2]$ in real life. Thus, agents enter the model as two-year-old children and exit at age 82.[10] This frequency allows us to investigate meaningful variations in school tracking ages. The structure implies that 20 generations are alive at every point in time. As in Lee and Seshadri (2019), we assume a unit mass of individuals in each period.

A life cycle can be structured into several stages, as illustrated in Figure 1: During the first four periods, a child lives with her parent, goes to school, and accumulates child skills. School tracking happens at the beginning of child period $j = 3$. At the beginning of child period $j = 5$, at age 18, the child becomes an independent adult, her child skills are transformed into adult human capital, and she can decide to go to college. Both college and non-college-educated types of labor are used, next to capital, by a representative firm to produce the final consumption good. Adult agents decide how much labor to supply until

[10] We choose this perhaps unorthodox timing to capture that in Germany, children are ten years old when parents make the secondary school track decision, which resembles reality in Germany. Appendix C gives an overview of the German Education System.

Figure 1: Timeline of Life-cycle Events



they retire at the beginning of $j = 17$, at age 66. During the working periods, human capital grows stochastically. Finally, in $j = 9$, when they are 34 years old, adults become parents of a child. Adults make inter-vivos transfers to their children when they turn 18 and become independent.[11]

2.1 Child Skill Formation

Every new child has an initial ability or human capital (skill) endowment, ϕ, which is imperfectly transmitted from her parent.[12] When children enter primary school, at the beginning of $j = 2$, the initial ability translates into a first child school skill level, θ_2, expressed in logarithm, which we refer to as child skills henceforth.

[11] For the remainder of the text, we will denote all child variables with primes whenever both parental and child states are present. Since an adult becomes a parent at age 32, the child of a parent who is in period j is in period $j' = j - 8$.

[12] As in Cunha and Heckman (2007), we do not differentiate between abilities and skills, as both are partly endogenously produced and partly exogenously determined pre-birth. The initial ability thus captures genetic components and investments made by parents into their child's development during early childhood, infancy, and even in-utero.

$$\theta_2 = \log \phi. \tag{1}$$

We think of child skills as stage-specific competencies during the schooling periods, $j = 2$ to $j = 5$, that can be observed by everyone and are rewarded on the labor market for both college and non-college-educated workers.

Subsequently, the evolution of skills depends on the schooling system. During primary school ($j = 2$), the system is *comprehensive*, meaning that there is only one track to which all schools belong, denoted by $S = C$. During secondary school, there are two distinct school tracks, a *vocational* track $S = V$ and an *academic* track $S = A$.[13] School tracks can differ in their *pace of instruction*, denoted by P^S, which reflects the differences in the intensity and depth with which school subjects are taught.[14] Notably, the pace of instruction in each school track is *endogenous* in the sense that the education policymaker can choose it in every period to achieve her goals. For our analysis, we assume that the policymaker has an efficiency goal and maximizes aggregate end-of-school skills.[15] We further assume that all classrooms and schools in the same track are identical. Thus, if a child is allocated to a particular track, we can think of her as attending a "representative" classroom and school

[13]While in principle a larger number of school tracks is conceivable, we restrict our analysis of tracking to two school tracks as this corresponds to a typical number across OECD countries. The two tracks typically serve the purpose of preparing children for academic higher education at a college or similar institution or to prepare children for a more vocational career.

[14]In Germany, the curricula and core subjects are no longer materially different across school tracks. The main difference between academic and vocational schools is that the former results in direct qualification to enter university, while the latter does not. In academic track schools, topics are generally taught more densely and comprehensively than in vocational track schools, preparing students for higher education. Moreover, students typically have more options for elective subjects at later stages of secondary school. Vocational track schools, by contrast, are less demanding in terms of the required learning effort, and graduation occurs after fewer years. A detailed comparison between the teaching intensity and learning goals across Germany is provided in Dustmann et al. (2017). Note that heterogeneity in instruction paces across tracks does not entail systematic differences in teacher quality or resources devoted to teaching across tracks that could also affect child skill development. In any case, the literature on international differences in student achievement tends to find limited effects of resources spent per student on learning outcomes (Woessmann, 2016). In Appendix C, we summarize information on expenditure per student as well as teacher quality across different school tracks in Germany.

[15]For example, in Germany, the curricula in the different tracks are set by each federal state under some general federal education goals. They consist of learning and competence goals, methods, and specific topics that should be taught in each school track, subject, and grade. The curricula are subject to frequent review and renewal. For example, as of 2023, 14 out of 16 federal states in Germany updated the curriculum in the last four years and 7 out of 16 in the last two years. The concrete implementation of the curricula, however, is in the hands of the teachers and individual schools, who have some discretionary margin to adjust the instruction paces to the needs of their pupils. We, therefore, view the pace-setting process as a mix of overarching learning goals and individual adjustment across school tracks. To the best of our knowledge, there is no clear teaching goal about the distribution of end-of-school skills formulated by German education policymakers.

for that track. This implies that all children in a given track are exposed to the same set of classroom and school peers.

The technology of (log) skill formation during the school years $j = 2, 3, 4$, of a child in school track S, is then given recursively by:

$$\begin{aligned} &\theta_{j+1} = \kappa\theta_j + \alpha\bar{\theta}_j^S + g(\theta_j, P_j^S) + \zeta E + \eta_{j+1} \\ &\eta_{j+1} \sim \mathcal{N}(0, \sigma^2_{\eta_{j+1}}). \end{aligned} \tag{2}$$

Next period's child's skills are directly affected by past skills θ_j and parental education E, which we take as a proxy for the home environment in which a child grows up, including differences in parental investments into child skills by parental background (Heckman and Mosso, 2014). By η_{j+1}, we denote unobserved i.i.d. shocks to the skills. This type of uncertainty in the formation of child skills is crucial for analyzing school tracking policies. We interpret these shocks as stemming, for example, from unexpected heterogeneity in child development speeds (such as late-bloomers), but also health shocks that can permanently influence the skill formation trajectory of a child.[16]

The school track can affect the next period's skills in two ways: First, through direct interactions with peers in a track, which affects future skills linearly through the average skill level of other children in school track S, denoted by $\bar{\theta}_j^S$, as is common in the peer effects literature (Sacerdote, 2011).[17] Second, through the pace of instruction in her school track, P_j^S, as governed by the function g. We assume this function takes the following form for a general instruction pace P_j in period j:

$$g(\theta_j, P_j) = \beta P_j + \gamma\theta_j P_j - \frac{\delta}{2} P_j^2. \tag{3}$$

This functional form implies firstly that for each skill level θ_j, there exists an individually-

[16]Our assumption of shocks as the source of child skill formation uncertainty is slightly different from the idea that the "true" academic potential of a child cannot be perfectly observed and must be learned over time from signals, such as school grades. We discuss the differences that would imply a model with imperfectly observed child skills in Appendix Section H.

[17]We concentrate on the case with a linear-only direct peer externality governed by α. As summarized in Epple and Romano (2011), many studies find that such linear-in-means peer effects are sizable and robust across settings. Evidence on non-linear peer effects in the classroom is more ambiguous. For that reason, we do not incorporate non-linearities in peer effects directly. Instead, we consider the endogenous setting of instruction levels across school tracks as a channel through which non-linear peer effects arise. We note, however, that non-linear peer effects could have important implications for the assessment of tracking policies. Moreover, we abstract from peer effects that operate through friends and a child's network outside of schools (see Agostinelli et al., 2023), as our data does not contain information on friendships.

optimal instruction pace, $P^*(\theta_j)$, that maximizes future skills. Secondly, if $\gamma > 0$, there is a positive complementarity between the individually optimal pace and the individual skill level, such that higher-skilled children also prefer a higher pace of instruction. This is motivated by evidence on the heterogeneous effects of teaching or instructional practices depending on prior student achievement and, in particular, by evidence on "match" effects between teaching practices and classroom skill composition (see Duflo et al. (2011), Aucejo et al. (2022) and references therein). As we will demonstrate theoretically in Section 3 and quantitatively in Section 5, this complementarity plays a central role in providing the rationale behind any efficiency argument in favor of school tracking policies.

Given (3), it is clear that aggregate learning is maximal if every child is taught at her preferred instruction pace in every period. However, there is only one instruction pace per school track. Given this constraint, a policymaker seeking to maximize expected future skills would then set the pace in each track to the one that is optimal for a child with exactly the *average skill level* in that track, as summarized in Lemma 1.

Lemma 1. *The pace of instruction a policymaker would set in each school track to maximize expected skills in the next period is given by*

$$P_j^S = P_j^*(\bar{\theta}_j^S) = \frac{\beta + \gamma\bar{\theta}_j^S}{\delta} \tag{4}$$

where $\bar{\theta}_j^S$ is the average skill level of children in track S.

Proof. Follows from taking the first order condition of the conditional expected value $\mathbb{E}(\theta_{j+1}|S)$ in (2) with respect to P_j^S using (3) and under the i.i.d. assumption of η_{j+1} and the fact that maximization of skills in each school track is a necessary condition for maximizing unconditional skills. □

According to Lemma 1, the instruction pace setting implies that future child skills depend non-linearly on her peers' skills. Specifically, skill gains decrease monotonically with the distance between a child's own skills and the average skill level in that track or equivalently with the distance between her optimal instruction pace and the one she is currently taught at.[18] Consequently, for a child with a low skill level, going to a school track with a high instruction pace tailored to a higher average skill level can be harmful to the point that

[18]See Appendix A for the derivation. Our formulation of learning hence implies that non-linear peer effects are driven by how the instruction levels are adjusted (as found, for example, in Duflo et al. (2011) or Lavy et al. (2012)). Moreover, it provides a micro-foundation for efficiency gains in average learning that stem from more homogeneous peer groups. We discuss the theoretical consequences of tracking under these assumptions in Section 3.

she actually learns less, despite being surrounded by better peers, than if she had attended a school with a lower pace. The average skill level in each track is an equilibrium object that depends, among other things, on parental skills, aspirations, and the track-specific labor-market outcomes that parents anticipate.

After finishing school, at the beginning of $j = 5$, child skills are transformed one-to-one into the first adult human capital level, h_5,

$$h_5 = \exp(\theta_5). \tag{5}$$

2.2 Preferences

We assume that the preferences over consumption and labor supply of adults in each period take the following form:

$$u(c_j, n_j) = \frac{(c_j/q)^{1-\sigma}}{1-\sigma} - b\frac{n_j^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} \tag{6}$$

where c_j denotes household consumption and q is an adult consumption-equivalent scale that is larger than one whenever there is a child in the household and one otherwise. Risk aversion is captured by σ. Individuals incur disutility from working n_j hours, which is governed by b and the Frisch elasticity of labor supply, γ. In each period, the maximum hours worked are normalized to 1.

Parents are altruistic as they take into account the utility of their child when making inter-vivos transfers. The strength of altruism is governed by a factor Λ. All future values are discounted by β.

2.3 Educational Choices

There are two types of educational choices agents make during their life. The first and novel education choice is the secondary school track parents choose for their children. We assume that the utility of parents also depends on the track their children attend, through a stochastic academic-track utility cost $\chi(E) \sim H^E(\chi)$, whose distribution can depend on parental education E. This will allow us to capture that empirically, the school track decision is significantly affected by parental socio-economic status, even conditional on school performance, test scores prior to the track decision, and the track recommended by primary school teachers. Moreover, when parents deviate from the primary school teacher's recommendation, it is usually toward their own education path.[19]

There may be multiple reasons behind these parent-specific academic track costs. For example, there may be a cost associated with acquiring information about school tracks that is lower whenever a parent went to that track herself. Similarly, parents may feel better able to support their child in a track they are more familiar with. Parents may also systematically over- or underestimate their children's potential or have strong preferences for their child following in their footsteps. Whatever their exact reason, deviations in parent's track choice from the recommended path may lead to the misallocation of children across tracks. For example, a child with low skills could be sent to the academic track by parents who have preferences for this track. This would lead to learning losses not only for the individual child but also create an externality for all other children as the instruction pace is endogenous to the peer composition.

Second, after finishing school, newly independent adults decide whether to go to college. In line with the literature (e.g. Daruich, 2022; Fuchs-Schündeln et al., 2022), we assume that going to college entails a "psychic" utility cost $\psi(S, \theta_5, \nu(E^p))$ that may depend on the secondary school track S, the end-of-school skills θ_5 and an idiosyncratic college taste shock, $\nu(E^p) \sim G^{E^p}(\nu)$, whose distribution may be influenced by the parent's education level E^p.[20]

This formulation can accommodate two important features of the transition between secondary and college education in the data. Firstly, the share of children with an academic track secondary school degree who get a college degree is significantly higher than those with a vocational secondary school degree.[21] Secondly, independently of the school track, the likelihood of college education in the data is increasing in the end-of-school skills.[22] Finally,

[19]See Appendix F for some reduced-form evidence on the school track choice and deviation by parental background. Importantly, children who deviate from the recommended school track perform differently than the others. Children who deviate from vocational to academic perform worse than the average kid in the academic track, and the reverse happens for children who deviate from academic to vocational. The fact that deviating from the recommended track does not seem to benefit children in terms of their achievements indicates that it is not the case that parents "know" the true potential of their child better or can support them better.

[20]We add the superscript "p" here to indicate that E^p is the college education of the parent of a newly independent adult who chooses her own college education E.

[21]In Germany, every graduate from an academic track secondary school gets an official qualification that allows for access to academic higher education institutions, while graduates from vocational tracks do not. To go to college, these must either get a qualification through "evening schools" or may be allowed access to certain university degrees after obtaining a higher vocational degree or after having worked for a certain number of years.

[22]Net of the above-explained effect coming through the secondary school track graduation, this may partly be because, for many university degrees, admission is competitive and often even requires a specific end-of-school grade average ("numerus clausus"). Of course, it could also simply reflect the selection of higher-skilled school graduates into an academic career, where these (mostly cognitive) skills are more valuable.

the random taste shocks reflect heterogeneity in the higher education decision coming from parental background or channels outside of the model, as is common in this literature.

2.4 Adult Human Capital, Labor Income and Borrowing

During the working career ($j = 5$ to $j = 16$), following Yum (2023), human capital grows according to:

$$h_{j+1} = \gamma_{j,E}\; h_j\; \varepsilon_{j+1}, \quad \log \varepsilon_j \sim \mathcal{N}(0, \sigma_\epsilon^2) \tag{7}$$

where $\gamma_{j,E}$ are age- and education-specific deterministic growth rates and ε_{j+1} are market luck shocks, which we assume follows an i.i.d. normal distribution in logs, with zero mean and constant variance σ_ε^2, as in Huggett et al. (2011). Human capital remains constant after retirement. Gross labor income is then given by:

$$y_j = w_E\; h_j\; n_j \tag{8}$$

where w_E denotes the effective wage per unit of human capital paid to workers with higher education E.

Note that all prices, including w_E, implicitly depend on the distribution of agents in the economy, which we suppress for notational convenience. After retiring, each agent receives retirement benefits $\pi_j(h_{17}, E)$, which depend on the last education-specific human capital level before retirement.[23] Throughout their life, adult agents can save into a risk-free asset a, which pays a period interest rate r. As in Lee and Seshadri (2019), we assume that each agent's borrowing is constrained by the amount that can be 100% repaid in the next period using a government transfer g. Moreover, agents cannot borrow against the future income of their children. The per-period borrowing constraint can thus be written as

$$a_{j+1} \geq \frac{-g}{1+r}. \tag{9}$$

In the following, we provide a recursive formulation of the agent's decisions in each life cycle stage.

[23] As is common in the literature, we let benefits depend on human capital in this way to proxy for lifetime earnings, which form the basis of pension benefits in reality.

2.5 Recursive Formulation of Decisions

At the beginning of each adulthood period prior to retirement, individuals learn about their market luck shock realization and, in case they have a child, about the child skill shock realization. Based on this information, they decide on consumption (c_j), savings (a_{j+1}), and hours worked (n_j). In addition, there are two education choices: the school track of her child in period $j = 11$ by the parent and the college decision in period $j = 5$ by the newly independent adult. Finally, parents decide on the inter-vivos transfer in period $j = 13$. All decisions are subject to the human capital growth technology (7), the borrowing constraint (9), a time constraint $n_j \in [0, 1]$ and a period budget constraint,

$$c_j + a_{j+1} = y_j + (1 + r)a_j - T(y_j, a_j) \tag{10}$$

where labor income is defined as in (8) and $T(y_j, a_j)$ gives taxes net of transfers, which consist of labor income and capital taxes.

2.5.1 Parenthood (Age 34-50, periods $j = 9, ..., 13$)

Parent with a Young Child ($j = 9,\ 10$) The state space in these periods consists of the parent's education E, her human capital, h_j, and her assets a_j. Parents observe their child's initial ability ϕ at the start of the first period of the child's life, when she is two years old. The skill of children at the beginning of their second period, at age 6, is given by (1).

Future child skills $\theta_{j'+1}$ evolve according to (2) given the optimal pace of instruction as defined in Lemma 1. In particular, primary schools are comprehensive track schools, such that the evolution of a child's skills depends on the average skill level of all children in their cohort $\bar{\theta}_{j'=2}$. The problem of the parent can then be written as:

$$\begin{aligned} V_j(E, h_j, a_j, \phi, \theta_{j'}) = & \max_{c_j, a_{j+1}, n_j} \left\{ u(\frac{c_j}{q}, n_j) + \beta\, \mathbb{E}\, V_{j+1}(E, h_{j+1}, a_{j+1}, \phi, \theta_{j'+1}) \right\} \\ & \text{s.t. } \theta_{j'+1} = \kappa\theta_{j'} + \ \alpha\bar{\theta}^S_{j'} + \ g(\theta_{j'}, P^*_{j'}(\bar{\theta}_{j'})) + \ \zeta E + \eta_{j'+1} \\ & (7) - (10) \end{aligned} \tag{11}$$

where expectations are taken over child skill shocks ($\eta_{j'+1}$), market luck shocks (ε_{j+1}), and in period $j = 10$ also over school track taste shocks $\chi(E)$.

The School Track Decision ($j = 11$) When the child turns ten, at the beginning of her third period of life, the parent decides on whether to send her child to the vocational or academic track school, $S \in \{V, A\}$. The decision of parents is not constrained by any education policy (but parents do generally obtain a track recommendation from their children's primary school). [24] Once a child is tracked, she remains in that track for two periods, until the end of secondary school, when she turns 18.[25] Parents make the track decision by comparing the value of sending the child to a vocational track school ($S = V$) with that of sending her to an academic track school ($S = A$). These (interim) values are given by

$$
\begin{aligned}
W_{11}(E, h_{11}, a_{11}, \phi, \theta_3, S) = \max_{c_{11}, a_{12}, n_{11}} & \left\{ u(\frac{c_{11}}{q}, n_{11}) + \beta \, \mathbb{E} \, V_{12}(E, h_{12}, a_{12}, \phi, \theta_4, S) \right\} \\
\text{s.t. } & \theta_4 = \kappa\theta_3 + \ \alpha\bar{\theta}_3^S + \ g(\theta_3, P_3^*(\bar{\theta}_3^S)) + \ \zeta E + \eta_4 \\
& (7) - (10)
\end{aligned}
\tag{12}
$$

for each track S. They encode several incentives that influence the track decision. On the one hand, academic track attendance makes, ceteris paribus, college access more likely, which results in higher human capital growth and productivity over the life cycle. Of course, the returns to college education depend on the demand for college-type labor. On the other hand, parents know that her child's skill formation depends on the average skill level in a school track $\bar{\theta}_3^S$, both directly through peer interactions but also indirectly through the endogenous optimal instruction pace P_3^S. Thus, parents need to anticipate the distribution of children across tracks when making the track decision, which becomes an aggregate state, which we keep implicit.

On top of that, the track decision is also affected by the stochastic academic track utility shock, $\chi(E) \sim H^E(\chi)$. Parents then make the discrete track choice using (12) after observing a draw of $\chi(E)$. Thus, we can define the value of a parent after this shock realization at the beginning of period $j = 11$ as

[24]This has become common practice in Germany, where in the majority of federal states, parents are completely free in making the secondary school track choice for their children. Only in three states, Bavaria, Thuringia, and Brandenburg, academic school track access is conditional on a recommendation by the primary school teachers. These recommendations are often tied to achieving a certain grade point average in Math and German in primary school. However, even in these states, children without a recommendation can take advantage of a trial period in an academic track school, after which the child will be assessed again.

[25]We abstract from track switches during secondary school, as these are relatively rare in the data. For example, in 2010/11, only around 2.5% of children in the first stage of secondary school in Germany switched school tracks (Bellenberg and Forell, 2012). Moreover, this number includes switches among different tracks that we group into the vocational track, so it is likely an upper bound of the track switches between the vocational and academic tracks. However, this does not preclude track switches between the end of secondary school and the beginning of possible tertiary education, which we allow in our model.

$$V_{11}(E,h_{11},a_{11},\phi,\theta_3) = \max_{S\in\{V,A\}} \{W_{11}(E,h_{11},a_{11},\phi,\theta_3,S=V), \\ W_{11}(E,h_{11},a_{11},\phi,\theta_3,S=A) - \chi(E)\}. \tag{13}$$

Remaining Parenthood ($j = 12, 13$) In period $j = 12$, when the child is 14 years old and starts the second period of secondary school, the parent solves the following problem:

$$\begin{aligned} W_{12}(E,h_{12},a_{12},\phi,\theta_4,S) = & \max_{c_{12},a_{13},n_{12}} \left\{ u(\frac{c_{12}}{q}, n_{12}) + \beta\, \mathbb{E}\, V_{13}(E,h_{13},a_{13},\phi,\theta_5,S) \right\} \\ & \text{s.t. } \theta_5 = \kappa\theta_4 + \alpha\bar{\theta}_4^S + g(\theta_4, P_4^*(\bar{\theta}_4^S)) + \zeta E + \eta_5 \\ & (7) - (10) \end{aligned} \tag{14}$$

where the child's school track S, which has been decided in the previous period, is now included in the parent's state space.

Just before her child reaches the age of 18 and becomes independent, the parent decides on a financial inter-vivos transfer that her child receives, a'_5, while taking into account the child's future value $V_{j'=5}$. As in Daruich (2022), we model this as an interim decision problem and assume that the parent already knows the realization of her market luck shock and her child's final skill shock but does not know the realization of the college taste shock $\nu'(E) \sim G^E(\nu')$. The transfer cannot be negative, so parents cannot borrow against the future income of their child. The value at the beginning of period 13 is then

$$V_{13}(E,h_{13},a_{13},\phi,\theta_5,S) = \max_{a'_5\geq 0} \left\{ \widetilde{V}_{13}(E,h_{13},a_{13}-a'_5) + \Lambda\, \mathbb{E}\, V_{j'=5}(\theta_5,a'_5,\phi,S,E,\nu'(E)) \right\} \tag{15}$$

where $\widetilde{V}_{13}$ is the value for a parent with savings a_{13} after the inter-vivos transfer has been made

$$\begin{aligned} \widetilde{V}_{13}(E,h_{13},a_{13}) = & \max_{c_{13},a_{14},n_{13}} \{u(c_{13},n_{13}) + \beta\, \mathbb{E}\, V_{14}(E,h_{14},a_{14})\} \\ \text{s.t. } & c_{13} + a_{14} + a'_5 = y_{13} + (1+r)a_{13} - T(y_{13},a_{13}) \\ & (7) - (9) \end{aligned} \tag{16}$$

so that the transfer a'_5 enters the budget constraint.

2.5.2 Work Life Without a Dependent Child (Age 18-34 and 50-66, periods $j = 5, 6, 7, 8$ and $j = 14, 15, 16$)

Independence ($j = 5$) At the beginning of adulthood, when the child turns 18, the state space of a newly independent adult comprises the secondary school track she graduated from S, end-of-school skills θ_5, initial assets a_5, which she received from her parents, initial ability ϕ and her parent's education E^p, which affects the distribution of the stochastic college taste shock $\nu(E^p)$. Conditional on the realization of that shock, the young adult first decides whether to go to college ($E = 1$) or not ($E = 0$) by solving the following problem:

$$\begin{aligned} V_5(\theta_5, a_5, \phi, S, E^p) = \max_{E \in \{0,1\}} \{ & W_5(E = 0, h_5, a_5, \phi), \\ & W_5(E = 1, h_5, a_5, \phi) - \psi(S, \theta_5, \nu(E^p))\} \end{aligned} \tag{17}$$

where W_5 denotes the values of college and non-college education, given by

$$\begin{aligned} W_5(E, h_5, a_5, \phi) = & \max_{c_5, a_6, n_5 \in [0, \bar{n}(E)]} \{u(c_5, n_5) + \beta \, \mathbb{E} \, V_6(E, h_6, a_6, \phi)\} \\ & \text{s.t. } (7) - (10) \end{aligned} \tag{18}$$

and end-of-school skills are transformed into adult human capital h_5 according to (5). Recall that the psychic utility cost of going to college $\psi(S, \theta_5, \nu(E^p))$ also depends on the secondary school track and end-of-school skills. While agents can work during college education, they only receive the vocational wage rate w_0. Moreover, obtaining a college education reduces the time available for work, as individuals spend part of their total time endowment studying, thus $\bar{n}(E = 1) < 1$.

Remaining Work Life ($6, 7, 8$ and $j = 14, 15, 16$) In periods 6 and 7, which correspond to ages 22 to 30, adults solve:

$$\begin{aligned} V_j(E, h_j, a_j, \phi) = & \max_{c_j, a_{j+1}, n_j} \{u(c_j, n_j) + \beta \, \mathbb{E} \, V_{j+1}(E, h_{j+1}, a_{j+1}, \phi)\} \\ & \text{s.t. } (7) - (10) \end{aligned} \tag{19}$$

In period $j = 8$, when they are age 30 to 34, adults know that they will have a child at the start of the next period. For that reason, they take expectations over the initial ability of their future child, ϕ', on top of the expectations over the market luck shock. Thus, we obtain that in period 8

$$
\begin{aligned}
V_8(E, h_8, a_8, \phi) = \max_{c_8, a_9, n_9} &\{u(c_8, n_8) + \beta\, \mathbb{E}\, V_9(E, h_9, a_9, \phi')\} \\
\text{s.t. } \log \phi' &= \rho_\phi \log \phi + \epsilon_\phi, \qquad \epsilon_\phi \sim \mathcal{N}(0, \sigma_\phi^2) \\
&(7) - (10)
\end{aligned} \tag{20}
$$

where ϵ_ϕ is an intergenerational shock. For periods $j = 14, 15, 16$, when they are age 54 to 66, adults are again without child and solve the standard life-cycle savings problem:

$$
\begin{aligned}
V_j(E, h_j, a_j) = \max_{c_j, a_{j+1}, n_j} &\{u(c_j, n_j) + \beta\, \mathbb{E}\, V_{j+1}(E, h_{j+1}, a_{j+1})\} \\
\text{s.t. } &(7) - (10)
\end{aligned} \tag{21}
$$

where the ability ϕ has already been transmitted to the child and does not enter the state space anymore. In the last period prior to retirement, $j = 16$, agents no longer need to take expectations over market luck shocks, as human capital remains constant during retirement.

2.5.3 Retirement, $j = 17, 18, 19, 20$

Everybody retires at the beginning of model period 17, corresponding to age 66, and receives retirement benefits $\pi_j(h_{17}, E)$. After period 20, at age 82, agents die with certainty and exit the model. The values in these periods are

$$
\begin{aligned}
V_j(E, h_{17}, a_j) = \max_{c_j > 0, a_{j+1}} &\{u(c_j, 0) + \beta V_{j+1}(E, h_{17}, a_{j+1})\} \\
\text{s.t. } c_j + a_{j+1} &= \pi_j(h_{17}, E) + (1 + r) a_j - T(0, a_j) \\
&\text{and } (9).
\end{aligned} \tag{22}
$$

2.6 Aggregate Production, and Government

A representative firm produces output according to the Cobb-Douglas production function $Y = AK^\alpha H^{1-\alpha}$, where A denotes total factor productivity, K is the aggregate physical capital stock, and H is human capital defined by:

$$
H = [\varphi H_0^{\sigma_f} + (1 - \varphi) H_1^{\sigma_f}]^{\frac{1}{\epsilon}}. \tag{23}
$$

Here, H_0 is the aggregate labor supply in efficiency units of workers with vocational higher education, and H_1 is that of workers with a college education. The physical capital stock depreciates at rate δ_f.

The government taxes labor income progressively, such that labor income net of taxes amounts to $y_{net} = \lambda y^{1-\tau_n}$ (Heathcote et al., 2017). It also taxes capital income linearly according to $\tau_a r a_j$. All tax revenue is used to finance retirement benefits π_j and fixed lump-sum social welfare benefits g that are paid to every household.

2.7 Equilibrium

We solve for the model's stationary equilibrium and its associated distribution using the numerical strategy in Lee and Seshadri (2019). Stationarity implies that the cross-sectional distribution over all states in every period j is constant across cohorts. Our model economy consists of 20 overlapping generations or cohorts at each time. The equilibrium requires that households and firms make optimal choices according to their value functions and firm first-order conditions, respectively. Moreover, the aggregate prices for physical capital and both types of human capital r, w_0, and w_1 are competitively determined and move to clear all markets. Note that we do not require the government budget to clear as well. Instead, we assume that all government revenues that exceed the financing of all social welfare programs result in wasteful government spending (or spending that is linearly separable in the utility of households).

A special feature of our model is that learning during the school years depends on the distribution of children across school tracks. Importantly, an equilibrium therefore requires that parents form expectations over the skill distribution across school tracks, which have to coincide with the actual distributions in equilibrium. Appendix B gives a detailed definition of the equilibrium.

3 Developing Intuition: School Tracking and Child Skill Formation

Our formulation of the child skill formation technology during the schooling years in (2) constitutes the novel cornerstone of our model. We now develop some intuition what it implies for skill accumulation with and without school tracking. Our focus in this section is exclusively on the secondary schooling years (periods 3 and 4), and we ignore transitions to higher education and the labor market. Moreover, we simplify parents' preferences, such that they only care about their child's expected end-of-school skills and have no other preferences regarding the school track choice. Finally, we assume for simplicity that $\kappa = 1$ and that there are no direct parental influences, $\zeta = 0$, nor stochastic parental track preferences, $\chi = 0$.

All other assumptions are maintained. In particular, policymakers set the instruction paces in each school track to maximize expected end-of-school skills, such that the pace-setting rule in Lemma 1 holds. Moreover, we assume that the distribution of child skills at the beginning of secondary school is normal and centered around 0.

3.1 Comprehensive School versus Tracking

We start by comparing a comprehensive schooling system (C), in which all children attend the same school track, to a tracking system (T) in which all children are tracked into a vocational or academic track. For simplicity, we consider only one period of schooling here. Thus, if θ_3 are the skills at the beginning of secondary school, θ_4 can be considered the skills at the end of school. A key implication of this simplifying assumption, when combined with the timing of skill shocks in (2), is that skill evolution during secondary school occurs as if there were no shocks to student's skill during that time. As we will see, this implies that aggregate end-of-school skills are always greater with tracking than in a comprehensive system.

The Allocation of Children across Tracks

To that end, we consider two alternative allocation mechanisms. In the first one, a policymaker (or a teacher) allocates children across tracks directly. As before, the goal of the policymaker is to maximize the expected end-of-school skills across all children ($\max_S \mathbb{E}(\theta_4)$).

The second alternative consists of each parent making the track decision unilaterally for her child i with skill level $\theta_{i,3}$. A parent's only goal is to maximize her child's expected end-of-school skill level ($\max_S \mathbb{E}(\theta_{i,4})$). Parents know the distribution of θ_3. We can thus think of this mechanism as a simultaneous move game played among parents, where each parent's strategy set consists of the two tracks she can send her child to, and the next period's skills give the payoffs.

Proposition 1 shows that, in both alternatives, the track decision that results in the optimum or equilibrium is governed by a sharp cut-off skill level. A policymaker would optimally split the distribution exactly at its mean.[26] Intuitively, this generates the highest aggregate end-of-school skills as it minimizes the variance of skills in each track, or in other words,

[26] A similar argument has been made repeatedly in the theoretical literature. See for instance, Epple and Romano (2011).

it creates peer groups that are as homogeneous as possible. In doing so, the policymaker internalizes that any effects coming from the direct peer externality offset each other across tracks. Thus, all gains achieved from making average peer skills in one track higher are lost as the average level in the other track becomes smaller.

In contrast, if parents are the decision-makers, they decide irrespectively of the aggregate outcomes. The equilibrium of this implied game still features a sharp skill threshold, which is characterized by the skill level at which a child's expected end-of-school skills are exactly equal in both tracks. This threshold is smaller than the optimal threshold a policymaker would pick whenever the direct peer effects are positive ($\alpha > 0$). The reason is that, because of positive direct peer effects, children with skills just below the policymaker's threshold would benefit individually from going to the academic track (with higher average skills). As parents do not internalize the effect of their decision on average skills in each track, they will therefore send their children to the academic track.

Proposition 1. *The allocation of children across tracks is characterized by a skill threshold $\tilde{\theta}_3$, such that all children with initial skills below $\tilde{\theta}_3$ go to one track and all children with initials skills above $\tilde{\theta}_3$ go the other track.*

- *If the policymaker does the track allocation, the optimal skill threshold corresponds to the average initial skill level $\tilde{\theta}_3^* = \mathbb{E}[\theta_3] = 0$.*
- *If parents do the track allocation, the skill threshold depends on the direct peer externality α. With $\alpha > 0$, the threshold is smaller than $\tilde{\theta}_3^*$.*[27]

Proof. In Appendix A. □

Next, we compare the comprehensive and tracking system in terms of their effects on end-of-school skills. We refer to an optimal tracking system, when the policymaker makes the track allocation with the goal to maximize end-of-school skills, as in Proposition 1.

The End-of-School Distribution

Provided that $\gamma \neq 0$ and $\delta > 0$, Proposition 2 shows that independently of the sorting mechanism, expected end-of-school skills in a full optimal tracking system are always larger than in a comprehensive system. Intuitively, this advantage comes from more homogeneous

[27] We rule out an (uninteresting) equilibrium of the track choice game in which parents randomly allocate their child into one of the two tracks, leading to the same distribution of skills in both tracks and, hence, the same pace of instruction.

peer groups in each track in terms of their skills. Since learning decreases with the variance of skills among children in a track, more homogeneity on average increases end-of-school skills. Therefore, the gain from tracking increases the smaller the conditional variance of skills across tracks, as given in equation (24). This gain from tracking further increases in the complementarity between own skills and instruction pace, γ. The stronger the complementarity, the more it pays to stratify children by their skills. Moreover, the advantage increases in the variance of initial child skills $\sigma^2_{\theta_3}$ but decreases in δ, which ultimately governs the concavity of learning with respect to the instruction pace.

A full tracking system may lead to larger inequality in end-of-school skills. In particular, condition (25) states that the variance of end-of-school skill might be larger in a tracking system with positive peer externalities if tracking occurs at the optimal skill threshold. This is more likely to hold the larger the direct peer externality and the larger the ratio $\frac{\beta\gamma}{\delta}$.

Similarly, an optimal tracking system necessarily leaves some children worse off compared to a comprehensive system. These children have initial skills around the tracking threshold and would be closer to their optimal instruction pace in a comprehensive system. In an optimal tracking system with $\tilde{\theta}_3 = 0$, these children thus occupy the center of the distribution and would, given a choice, prefer a comprehensive system. If there are no direct peer effects, an equal share of children in both tracks lose relative to the comprehensive counterpart. However, with positive peer effects, the losses are concentrated among the track with the lower average peer level. This reflects a robust finding of the empirical school tracking literature that the children at the bottom of the skill distribution suffer from a tracking system (e.g. Matthewes, 2021).

Proposition 2.

- *Expected end-of-school skills in a full tracking system are larger than in a fully comprehensive system. This holds regardless of who makes the track decision, i.e., regardless of the tracking skill threshold $\tilde{\theta}_3$. The gain from tracking is given by*

$$\mathbb{E}(\theta_4|T) - \mathbb{E}(\theta_4|C) = \frac{\gamma^2}{2\delta}\left(\sigma^2_{\theta_3} - \mathbb{E}(Var[\theta_3|S])\right). \tag{24}$$

- *The end-of-school skill distribution in a full tracking system has a "fatter" right tail. In case of tracking at the optimal skill threshold $\tilde{\theta}_3 = \mathbb{E}(\theta_3)$, the variance of end-of-school skills in a full tracking system is larger than the variance in a fully comprehensive system iff*

$$\alpha^2 + 2\alpha\left(1 + \frac{\beta\gamma}{\delta}\right) - (8-\pi)\frac{\gamma^4}{\pi\delta^2}\sigma^2_{\theta_3} > 0. \tag{25}$$

- *Children with initial skills inside a non-empty interval lose from a full tracking system in terms of their end-of-school skills relative to a fully comprehensive system. With $\alpha = 0$, the losses are symmetric in both tracks. With $\alpha > 0$, the losses are concentrated in the track with the lower average skill level.*

Proof. In Appendix A. □

As already mentioned, the main reason why T always beats C here in terms of aggregate skills is the simplifying assumption that there are no skill shocks during students' time in school. As a result, the tracking decision made at the start of secondary school is optimal throughout secondary school.

3.2 Early versus Late Tracking

Let us now consider a two-period secondary schooling system, like in our full model, where there can be skill shocks during the time students are in secondary school. In this case, the skills at the end of school are θ_5. We are interested in a comparison between the end-of-school skill distribution in an early tracking system, ET, and a late tracking system, LT. In both cases, the allocation of children to tracks is done optimally by a policymaker, maximizing the expected aggregate end-of-school skills ($\max_S \mathbb{E}(\theta_5)$). The early tracking system is characterized by a track allocation in $j = 3$ that is maintained throughout secondary school. That is, there is no re-tracking. The late tracking system is characterized by all children going to a comprehensive school in the first period, followed by tracking at the beginning of the second secondary school period ($j = 4$). Hence, in the case of LT, children are allocated to school tracks after the skill shocks η_4, while in the case of ET, the school track decision was made before the realization of these skill shocks.

Proposition 3 shows that expected end-of-school skills in an optimal LT system can be larger than in an optimal ET system if the variance of the skill shocks is large enough. Intuitively, this represents the key disadvantage of early tracking. Since the first track allocation is maintained throughout secondary school, it does not correct for skills shocks during that time. As a result, some students are mismatched in the second period of secondary school ($j = 4$). The LT system avoids this mismatch by making the track allocation later. But this comes at the cost of less aggregate skill accumulation during the C stage. Hence, when students are subject to skill shocks during secondary school, there is a trade-off between the pace of learning in the first stage of secondary school and the quality of the student-track match in the second stage of secondary school.

Proposition 3. *Expected end-of-school skills in the two-period model are larger in an optimal late tracking system than in an optimal early tracking system iff*

$$\frac{\sigma^2_{\eta_4}}{\sigma^2_{\theta_3}} > 1 + \alpha + \alpha^2 + \beta + \frac{\beta^2}{2} + 2\alpha(1+\beta) + \frac{\gamma^2}{2\pi}\sigma^2_{\theta_3}. \tag{26}$$

Proof. In Appendix A. □

These results illustrate that the child skill technology alone entails non-trivial theoretical implications for the effects of school tracking policies on end-of-school skills. In particular, even when the track allocation is performed optimally, in the sense of efficiency-maximizing, the timing of tracking balances a trade-off between efficiency gains from learning in more homogeneous peer groups and those from the ability to react to child skill shock realization.

In sum, our parsimonious child skill technology can therefore accommodate the ambiguous empirical findings on the effects of tracking on the level of educational achievements, in addition to the estimated association of tracking with increases in inequality and disproportional disadvantages among the lower-skilled groups. However, this alone does not allow us to quantify the macroeconomic effects of school tracking policies. Indeed, the quantitative importance of these forces for *economic outcomes* within and across generations not only depends on the estimates of the child skill technology parameters and the size of the skill shock variances, but also on how they interact with other essential features of the model (and reality). For example, second-chance opportunities at the time of the college decision may make the effect of the (early) track choice less consequential for labor market outcomes. On the other hand, asymmetric parental utility costs of school tracks may reinforce inter-generational persistence of education, while harming learning efficiency during the school years. Finally, the track decision, in reality, is likely not just concerned with purely maximizing skills but takes into account future labor market prospects, which in general equilibrium, also depend on the share of children attending each track. To quantify these channels through the lens of our model, we now describe the calibration procedure.

4 Model Calibration

We calibrate the model to the German Education System (described in detail in Appendix C) following a two-step approach. In the first step, we estimate the parameters of the child skill formation technology during the school years, as well as other selected model parameters directly from the data. In the second step, the remaining parameters are estimated using the simulated method of moments by matching the moments from the stationary equilibrium distribution of the model to their empirical counterparts. Table 4 summarizes the externally calibrated parameters, and Table 5 presents the internally estimated ones.

4.1 Data and Sample Selection

All externally estimated parameters in the first step and moments used in the second step are based on two data sources. The first source is the German National Educational Panel Study (NEPS), which comprises detailed longitudinal data on the educational process, acquired competencies, as well as the learning environment, and context persons of six cohorts of participants in nationally representative samples in Germany, starting in 2010 (Blossfeld et al., 2019).[28] A key component of the information collected is regular standardized assessment tests of school children's competencies in areas such as mathematics, reading, sciences, vocabulary, or grammar, combined with specific wave weights.[29] In addition, there is information about school track recommendations and the final parental school track choices.We restrict the sample to individual observations containing information on the school and class a child attended in a given year.

The second data source is the German Socioeconomic Panel (SOEP), an annual representative survey from which we use the 2010-2018 waves (Goebel et al., 2019). The data contains rich information on labor supply, income, and education on the individual level. We use this data source primarily to construct empirical moments for the working stage of the life cycle, as will be detailed below. For this reason, the only sample selection we do is dropping those workers with hourly wages below the first and above the 99th percentile while keeping both workers and non-workers. We convert all income data to 2015 Euros using a CPI index for inflation adjustment.

We begin by detailing how we measure, identify, and estimate the parameters of the child skill formation technology, as these constitute the most critical ingredient of our model. Then, we describe the functional forms and estimation strategies for all remaining parameters.

[28] The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network. We use data from Starting Cohorts 2,3, and 4 (NEPS Network, 2022).

[29] See also Appendix Section D for more details on the tests as well as the scaling procedure adopted by the NEPS.

4.2 Estimation of the the Child Skill Formation Technology

We specify the empirical analog of the production technology of (the logarithm) of child i's skills that we take to the data as follows:[30]

$$\theta_{i,j+1} = \omega_{0,j} + \omega_{1,j}\theta_{i,j} + \omega_{2,j}\bar{\theta}^S_{-i,j} + \omega_{3,j}\theta^2_{i,j} + \omega_{4,j}(\theta_{i,j} - \bar{\theta}^S_j)^2 + \omega_{5,j}E_i + \eta_{i,j+1}, \tag{27}$$

Note that (27) is a rearranged version of the child skill technology (2) after substituting in (3) and the optimal pace of instruction in each school track as given by Lemma 1.[31] Moreover, in principle, we allow all child skill technology estimates to be specific to the period j.

In the estimation, we also distinguish between $\bar{\theta}^S_{-i,j}$, which denotes the average skill level of the child i's *classroom* peers, and $\bar{\theta}^S_j$, which refers to the average skill level of all children in a school that belongs to track S. Note that in the model, $\bar{\theta}^S_{-i,j} = \bar{\theta}^S_j$, since we assume a representative school and class per track (or alternatively, identical classes conditional on school tracks). In the data, however, there is heterogeneity across classes, even within schools and tracks. Since we are interested in capturing skill development effects that arise from direct interactions with peers, which are likely occurring in a specific classroom, we exploit that heterogeneity in the estimation.[32] Finally, the intercept $\omega_{0,j}$ can be a function of age and gender in the empirical estimation, and the parental educational attainment E is a time-constant dummy that equals one if child i comes from a household in which at least one parent is college educated.

As is common in the child skill formation literature (Cunha et al., 2010; Agostinelli and Wiswall, 2016), we think of child skills θ as latent variables that are only imperfectly measured in the data. Therefore, we employ a log-linear measurement system for latent skills,

[30]Following the work in Cunha et al. (2010), much of the empirical and quantitative literature using child skill formation technologies employs parametric specifications of the constant elasticity of substitution (CES) form. As noted in Agostinelli and Wiswall (2016), this requires, under standard parameter restrictions, that all input factors are static complements. An alternative is to use a nested CES structure as in Fuchs-Schündeln et al. (2023); Daruich (2022). To retain tractability, we follow Agostinelli and Wiswall (2016) and opt for the trans-log approach. In our formulation, all inputs into child skill formation, and in particular school inputs and parental inputs are therefore substitutes, which is in line with the literature (Kotera and Seshadri, 2017). We also experimented with relaxing this assumption by including interaction terms between school inputs and parental education, which were, however, insignificant.

[31]The coefficients $\omega_{n,j},\ n = 0, ..., 5$ relate to those in (2) and (3) as follows: $\omega_0 = \frac{\beta^2}{2\delta}, \omega_1 = (\kappa + \frac{\beta}{\gamma}\delta), \omega_2 = \alpha, \omega_3 = -\omega_4 = \frac{\gamma^2}{2\delta}$, and $\omega_5 = \zeta$ for all j. We formally test the restriction $\omega_3 = -\omega_4$ after the estimation.

[32]Given that we control for school fixed effects in the estimation, our identification of the direct peer effects is therefore close to the literature on estimating peer effects using classroom-fixed-effects methods within the same schools (see the discussion in Epple and Romano, 2011). This also has the added benefit that we can identify a model that includes $\bar{\theta}^S_{-i,j}, (\bar{\theta}^S_j)^2$, and the interaction $\theta\bar{\theta}^S_j$, even if we consolidate schools into a maximum of two school tracks in the data, which, as discussed in Appendix C resembles reality in Germany over the past decade.

using a series of achievement test scores as noisy measures of child skills in each period.[33] The identification strategy of the scales and loadings of each measure using their covariances follows Cunha et al. (2010). We aggregate the individual measures into a composite unbiased index using Bartlett factor scores, as in Agostinelli et al. (2023), to account for measurement error. Appendix E details skills measurement and the estimation procedure.

We present the estimates of the child skill production technology parameters in Table 1. Note that the estimates are based on the NEPS Starting Cohort 3 data, which follows children through secondary school. Since, before grade 5, children are in a unique school track, we cannot estimate the age-specific coefficients for period 2. In addition, in grade 12, some parts of the tests are track-specific, which makes the estimates unreliable for period 4. For those reasons, we assume that the estimates of the child skill technology parameters ω_2, ω_3, and ω_4 between school grades 5 and 9 are representative of the entire schooling career. That is, we drop the j index on those technology parameters.

Recall that $\theta_{i,j}$ is the logarithm of child skills. Hence, we can interpret the coefficients as elasticities. Thus, $\hat{\omega}_1 = 0.66$ means that a 1% increase in latent skills at the beginning of primary school is associated with a 0.66% increase in end-of-primary school skills. This own-skill productivity is close to the literature's common values (see estimates in Cunha et al. (2010); Agostinelli et al. (2019)). During secondary school, the estimated coefficient $\hat{\omega}_2$ is positive but rather small and statistically insignificant. Existing estimates of linear-in-means peer effects models range from small negative effects to large positive effects of a one unit increase in average peer test scores on student achievement.[34] Translating our estimates into such an effect, we find that a one unit increase in average peers' test scores raises own future tests by around 0.01. As such we are at the lower end of typical estimates during primary and secondary school, which is line with other research that uses within-school classroom variation (see Epple and Romano, 2011) that typically arrive at lower estimates compared to studies that use some form of random assignment of peers. Finally, the estimated coefficient $\hat{\omega}_4$ is negative and statistically significant at 10%. It indicates that a 1% increase in the squared distance to the average skill level in a track is associated with an up to 0.011% decrease in the next period's skills. This lends empirical support to the idea that the instruction pace in every track is tailored to the average skill level, and deviations,

[33] Regular independent tests of children's achievements in domains such as mathematics or languages are a key component of the NEPS data. We provide more information on the constructed test scores in Appendix D. As argued in Borghans et al. (2008), achievement test scores measure both cognitive and non-cognitive skills.

[34] Table 4.2. in Sacerdote (2011) provides an overview about existing estimates using a variety of identification strategies.

in both directions, from this level can hurt individual skill development. Importantly, we cannot reject the hypothesis that $\hat{\omega}_3 = -\hat{\omega}_4$ which is in line with our assumptions in Section 3. [35]

The final estimates we use to parameterize the child skill formation technology are ω_n for $n = 2, 4$ as reported in Table 1. We set $-\omega_3 = \omega_4 = \hat{\omega}_4$. The parameter $\omega_{1,3}$ and $\omega_{5,3}$ come from Table 1, while $\omega_{1,2}$, $\omega_{1,4}$, $\omega_{5,2}$, and $\omega_{5,4}$ are estimated internally to match the own-skill elasticities from a regression of future skills on past skills and parental education as reported in Table 2. Finally, the constant parameter ω_0 is set to zero.

Table 1: Child Skill Technology Parameters Estimates

Coefficient	Variable	Dependent Variable: $\theta_{i,j+1}$ Grade 9 on Grade 5
$\hat{\omega}_{1,3}$	$\theta_{i,j}$	0.664***
		(0.022)
$\hat{\omega}_2$	$\bar{\theta}_{-i,j,S}$	0.003
		(0.020)
$\hat{\omega}_3$	$\theta_{i,j}^2$	0.008*
		(0.004)
$\hat{\omega}_4$	$(\theta_{i,j} - \bar{\theta}_{j,S})^2$	-0.011*
		(0.006)
$\hat{\omega}_{5,3}$	$E = 1$	0.034***
		(0.010)
Obs.		1,847

Notes: This table presents the coefficients of regressions of skills in grade 9 on skills in grade 5, skills squared, the average skill level of peers, distance to the average skill in the track squared, and parent's education dummy. Standard errors are clustered at the school level. We control for year of birth, gender, and school-fixed effects. Source: NEPS.

[35] The estimated negative effect $\hat{\omega}_4$ is therefore conform with findings in the literature, which test the effects of skill-based tracking on later achievement directly (for example Duflo et al. (2011) argue that the large achievement gains of tracked students relative to non-tracked students are the result from indirect effects of peers that operate through the adjustment of teaching behavior) or test the effects of classroom heterogeneity on achievement. As summarized in Sacerdote (2011), many, but not all, findings in this literature point to the fact that classroom heterogeneity reduces test scores, which is consistent with the idea that tracking raises outcomes in both tracks. In terms of effect sizes, it is difficult to compare our estimates to existing ones as we are measuring the heterogeneity across tracks directly and not across classrooms.

Table 2: Evolution of Child Skills

	Grade $\theta_{i,j+1}$ on Grade $\theta_{i,j}$ and E		
Dependent Variable:	Grade 4 (Cohort 2)	Grade 9 (Cohort 3)	Grade 12 (Cohort 4)
Panel A: All students			
$\theta_{i,j}$	0.649***	0.811***	
	(0.011)	(0.016)	
$E = 1$	0.072***	0.044***	
	(0.007)	(0.009)	
Obs.	4,023	2,070	
Panel B: Academic students			
$\theta_{i,j}$	0.566***	0.745***	0.825***
	(0.019)	(0.025)	(0.019)
$E = 1$	0.049***	0.035***	0.033***
	(0.011)	(0.012)	(0.009)
Obs.	1,371	1,195	2,327

Notes: This table presents the coefficients of regressions of current skills on past skills and parents' education dummy. Standard errors are clustered at the school level. Models control for year of birth, gender, and school-fixed effects. Source: NEPS.

4.3 Remaining Parameters

4.3.1 Preferences

We set the inverse elasticity of intertemporal substitution to $\sigma = 2$, a value that is common in the literature. The Frisch elasticity of labor supply is set to 0.5. The disutility shifter b is estimated internally to match the average time worked in the SOEP data, which is 0.36 when the total time available after sleep and self-care is assumed to be 13 hours on a weekday and normalized to 1.

We internally calibrate the time discount factor β, so the equilibrium interest rate amounts to 4% annually. The altruism parameter Λ is calibrated such that the ratio of average inter-vivos transfers to average labor income in the model corresponds to average higher education costs of children to average four-year labor income in the data. According to a 2016 survey by the German Student Association, the monthly costs of living during the higher education stages for a student without children are, on average, 830 Euros per month (Dohmen et al., 2019). We expect the parents to bear the bulk of these costs and assume that they support their child for an average of four years (the length of time it takes on average to complete higher education studies). Then, the ratio of total costs to average 4-year labor income is approximately 0.49, which we take as our target moment.

4.3.2 Academic School Track Costs

The stochastic school track costs $\chi(E)$ are assumed to follow the distribution $\chi(E) \sim H^E(\chi) \equiv \mathcal{N}(\mu_{\chi,E}, \sigma_\chi^2)$.

We parameterize the mean $\mu_{\chi,E}$ as follows:

$$\mu_{\chi,E} = \mu_{\chi,A} + \begin{cases} \chi_1 & \text{if } E = 1 \\ \chi_0 & \text{if } E = 0, \end{cases} \tag{28}$$

so that $\mu_{\chi,A} > 0$ represents a uniform utility cost of academic-track attendance (for example, stemming from the academic track being more demanding and psychologically taxing), and the parameters χ_0 and χ_1 represent asymmetric preferences or costs for the academic track by parental college education. We calibrate χ_0 and χ_1 to match the share of deviations from secondary school track recommendations by parental education in the data, while $\mu_{\chi,A}$ is calibrated to match the overall share of academic track recommendations (0.44). Primary school teachers typically give these recommendations before the transition to secondary school. They are based on both a reflection of the child's achievement during primary school and the teachers' assessment of the academic potential and success probability of the child in an academic track school. Thus, we argue that the recommendations are forward-looking and, since primary school teachers typically observe the children over multiple years every day during the week, based on a similar information set as the parents. Therefore, we consider the recommended school track in the model as the one a parent would have chosen without any specific school track taste ($\chi_1 = \chi_0 = 0$). Then, deviations from that unbiased track choice by parental education map into deviations from teacher recommendation.[36] The variance of the track tastes σ^2_χ is calibrated to match the variance of the residuals coming from a regression of school track on end-of-primary-school skills, which is 0.166.

4.3.3 Initial Child Skills, and Child Skill Shocks

The transmission of innate ability ϕ, which equals the initial child skill level, across generations follows an AR(1) process with persistence coefficient ρ_ϕ and variance σ^2_ϕ. Since the innate ability is designed to capture any residual correlation in economic outcomes across generations, we calibrate it to match the intergenerational elasticity of incomes in Germany. Kyzyma and Groh-Samberg (2018) estimate an elasticity between the income rank of individual labor earnings between children and parents using the SOEP data of 0.24, which we take as our target statistic.[37] The variance σ^2_ϕ is then estimated to match the variance of pre-school skills in the data, which we normalize to 0.1.

An integral part of the child skill development is the presence of unforeseeable, permanent shocks to child skills. As discussed in Section 3, the size of such shocks has important implications for the effects of school tracking policies as they can give rise to efficiency losses from early tracking. To quantify the importance of child skill shocks in our model, we internally estimate the shock variance $\sigma^2_{\eta,j+1}$, for $j = 2, 3, 4$. As target moments, we choose the correlation of a child's skill percentile rank across periods. In this way, we capture all changes in a child's relative position in the skill distribution in a given period that cannot be accounted for by the skill formation technology or track choices.

[36] Details on these moments are given in Appendix G.

[37] As is common in the literature, Kyzyma and Groh-Samberg (2018) compute the correlation of income ranks using average labor earnings over five years. We compute rank-rank correlations of four-year labor income, according to the period length of our model, and then compare the ranks of 30-34-year-old children to those of their parents when they were 46 to 50 years old, which is similar to the sample used by Kyzyma and Groh-Samberg (2018). As in Lee and Seshadri (2019), we normalize average labor income across the entire working population to be one in the data and in the model. In the latter, we do this by setting the technology parameter A in the firm production function.

4.3.4 College Costs

We parameterize the "psychic" college cost function following Daruich (2022):

$$\begin{aligned} \psi(S, \theta_5, \nu(E^p)) &= \exp(\psi_0 + \psi_{S=V} + \psi_\theta \theta_5 + \nu(E^p)) \\ \nu(E^p) &\sim G^{E^p}(\nu) \equiv \mathcal{N}(\mu_{\nu,E^p}, \sigma_\nu^2). \end{aligned} \tag{29}$$

We estimate the two parameters ψ_0 and $\psi_{S=V}$ of the college costs to match the share of graduates from an academic secondary school who follow up with a college education and the share of vocational secondary school graduates that obtain a college education. We discipline the coefficient ψ_θ that multiplies end-of-school skills by matching the regression coefficient on test scores from a regression of a college graduation dummy on end-of-school test scores, controlling for the secondary school track.

We calibrate the two parental education-specific means of the college taste shock parameters to be symmetric deviations from 0, such that $\mu_{\nu,E^p=1} = \Delta(\mu_{\nu,E^p})$ and $\mu_{\nu,E^p=0} = -\Delta(\mu_{\nu,E^p})$ to match the ratio of the share of children from college-educated parents who themselves go to college (0.63) over the share of children from non-college-educated parents who go to college (0.20) in the data. Finally, we calibrate the variance of these shocks, σ_ν^2, to match the variance of the residuals from the above regression of college education on end-of-school skills and school track, which is 0.137.

The final component of college costs is not a part of the "psychic" costs but reflects the time cost of obtaining a college education. We assume that studying for a college degree takes away around 60% of the total time available for work for four years or one model period.[38] Thus, we set the maximum remaining time during the higher education stage to $\bar{n}(E = 1) = 0.40$.

4.3.5 Human Capital Growth

We estimate the deterministic human capital growth profiles for both types of education, $\{\gamma_{j,E}\}$, $j = 5, ..., 16$ using wage regressions in the SOEP data, following the approach in Lagakos et al. (2018).[39] The resulting experience-wage profiles for four-year experience bins are shown in Table 3, expressed in growth relative to the previous bin. We set the $\{\gamma_{j,E}\}_{j=5}^{16}$ parameters to these values.

[38] A standard estimate is that full-time studying takes around 40 hours per week, which amounts to around 60% of the maximum weekly work hours, which we set to 65. Moreover, the average study length in Germany is eight semesters or four years.

[39] Concretely we create, separately for each education group, four-year work experience bins. We then estimate Mincer regressions of wages on years of schooling and potential work experience, controlling for

Finally, we calibrate the variance of the market luck shocks, σ_{ε}^2, such that our model replicates the standard deviation of (normalized) labor income across the working-age population in the data, which is around 0.86.

Table 3: Human Capital Growth Profiles

Experience (Years)	Wage Growth Non-College	College
0	1.00	1.00
4	0.96	1.15
8	1.09	1.19
12	1.10	1.11
16	1.04	1.06
20	1.02	1.01
24	1.00	0.99
28	1.01	0.97
32	0.99	0.98
36	1.01	0.99
40	0.99	1.01

Notes: This table provides wage growth estimates by year of experience and educational attainment. Source: SOEP

4.3.6 Firms and Government

Following large parts of the literature, we set the capital share in the aggregate production function to $\alpha = 1/3$. Moreover, we set $\sigma_f = 1/3$ such that the elasticity of substitution between college and non-college human capital in the firm production equals 1.5 (Ciccone and Peri, 2005). The weight on non-college human capital in the CES aggregator, φ, is estimated internally. Following the arguments in Lee and Seshadri (2019), we calibrate it to match the share of college-educated workers in the SOEP data. The TFP parameter A is

time and cohort effects of the form:

$$\log w_{ict} = \alpha + \beta s_{ict} + \delta x_{ict} + \gamma_t + \zeta_c + \epsilon_{ict},$$

where w_{ict} is the wage of individual i, who belongs to birth cohort c and is observed at time t. Wages are defined as total annual labor earnings divided by hours worked. We denote by s_{ict} the years of schooling and by x_{ict} work experience, which is defined as

$$x_{ict} = age_{ict} - 18 \text{ if } s_{ict} < 12$$
$$x_{ict} = age_{ict} - s_{ict} - 6 \text{ else.}$$

We assume no experience effect on wage growth in the last eight years of work to disentangle time from cohort effects, following the HLT approach in Lagakos et al. (2018).

calibrated such that the model produces average earnings of 1.

Regarding the tax-related parameters, we set the labor income tax scale to $\lambda = 0.679$ and the labor tax progressivity parameter to $\tau_l = 0.128$ following estimates in Kindermann et al. (2020). The linear capital tax is set to $\tau_a = 0.25$, corresponding to the final withholding tax rate on realized capital gains, interest, and dividends in Germany. The size of the lump sum government transfers is set to $g = 0.06$, which in equilibrium amounts to 6% of average labor earnings. Finally, we set pension benefits to $\pi_j(h_{17}, E) = \Omega h_{17} w_E$ during retirement and calibrate the scale parameter Ω internally, such that the average replacement rate corresponds to 40% (Mahler and Yum, 2023).

Table 4: Parameters calibrated externally

Parameter	**Value**	**Description**	**Source**
Household			
σ	2.0	Inverse EIS	Lee and Seshadri (2019)
γ	0.5	Frisch Elasticity	Fuchs-Schündeln et al. (2022)
q	1.56	Equiv. Scale	Jang and Yum (2022)
$\bar{n}(E = 1)$	0.40	Time Cost of College	40 hours/week for 4 years
Firm			
σ_f	1/3	E.o.S (H_0, H_1)	Ciccone and Peri (2005)
δ_f	6%	Annual Depreciation	Kindermann et al. (2020)
Government			
τ_n	0.128	Labor Tax Progressivity	Kindermann et al. (2020)
λ	0.679	Labor Tax Scale	Kindermann et al. (2020)
τ_a	0.25	Capital Tax Rate	Tax Rate on Capital Gains in Germany
g	0.06	Lump-sum Transfers	6% of Annual Labor Income

Notes: This table presents the externally calibrated parameters and their corresponding sources.

4.4 Method of Simulated Moments Estimation Results

In total, we calibrate 26 parameters internally using the method of simulated moments to match 26 target data moments. The parameters, their estimated values, model-implied moments, and target data moments are presented in Table 5.

The model fits the data well, both in terms of aggregate moments and concerning the distribution of child skills, school tracks, and higher education. For example, the share of college graduates in the simulated economy is 35%, which is in line with the German data in the 2010s. The share of children in an academic track school is 44%. The model also matches the transition rates from academic and vocational secondary school into college (at around 66% and 11%).

Table 5: Internally Calibrated Parameters

Parameter	Value	Description	Target	Data	Model
Preferences					
β	0.935	Discount Factor	Annl. Interest Rate	0.04	0.04
b	20.7	Labor Disutility	Avrg. Labor Supply	0.36	0.36
Λ	0.31	Parental Altruism	Transfer/Income	0.49	0.49
School Track Tastes					
$\mu_{\chi,A}$	0.048	Uniform A-Track Costs	Share A-Track Recommend.	0.44	0.44
χ_0	0.0020	Mean A-Track Cost if $E = 0$	Share of Dev. from A if E = 0	0.16	0.16
χ_1	-0.0036	Mean A-Track Cost if $E = 1$	Share of Dev. from V if E = 1	0.23	0.23
σ_χ	$0.17 \cdot 10^{-3}$	Std. A-Track Cost Shock	Reg. S on θ: var(residuals)	0.166	0.168
Child Skill Technology					
$\omega_{1,2}$	0.65	Own Skill Elasticity ($j = 2$)	Reg. θ_3 on θ_2 & E: coef. θ_2	0.649	0.649
$\omega_{5,2}$	0.072	Coefficient on E ($j = 2$)	Reg. θ_3 on θ_2 & E: coef. E	0.072	0.072
$\omega_{1,4}$	0.81	Own Skill Elasticity ($j = 4$)	$S = 1$, Reg. θ_5 on θ_4 & E: coef. θ_4	0.825	0.812
$\omega_{5,4}$	0.032	Coefficient on E ($j = 4$)	$S = 1$, Reg. θ_5 on θ_4 & E: coef. E	0.033	0.032
Initial Skills and Ability Transmission					
σ_ϕ	0.032	Std. of Ability	Variance initial skills	0.10	0.12
ρ_ϕ	0.9	Persistence of Ability	IGE (income rank)	0.24	0.23
College Costs					
ψ	0.77	Intercept	Share in CL from A-Track	0.66	0.65
ψ_V	0.16	Add. Costs for V-Track	Share in CL from V-Track	0.11	0.11
ψ_θ	0.35	Coefficient on θ_5	Share in CL from non CL HH	0.20	0.28
$\Delta(\mu_{\nu,E^p})$	0.034	Diff. in Means by E^p	Share in CL from CL HH	0.63	0.49
σ_ν	0.008	Std. Taste Shock	Reg. E on θ & S: var(residuals)	0.137	0.138
Idiosyncratic Shocks					
σ_ε	0.011	Std. Luck Shock	Std(Log Labor Income)	0.86	0.84
σ_{η_3}	0.052	Std. Learning Shock $j = 3$	$\text{Rank}_{j=2}$-$\text{Rank}_{j=3}$	0.72	0.73
σ_{η_4}	0.030	Std. Learning Shock $j = 4$	$\text{Rank}_{j=3}$-$\text{Rank}_{j=4}$	0.79	0.80
σ_{η_5}	0.032	Std. Learning Shock $j = 5$	$\text{Rank}_{j=4}$-$\text{Rank}_{j=5}$ if $S = 1$	0.74	0.75
Miscellaneous					
Ω	0.1	Pension Anchor	Replacement Rate	0.40	0.40
A	3.31	TFP	Avrg. Labor Earnings	1.0	1.0
φ	0.543	Weight V. Human Capital	College Share	0.35	0.35

Notes: This table presents the internally calibrated parameters, targeted moments, and their model-generated counterfactuals.

Parental school track utility costs significantly affect the school track decision, both in the model and the data. In particular, around 23% of college-educated parents overrule a recommendation for their child to go to a vocational track school, while 16% of non-college parents overrule an academic track recommendation in favor of a vocational track school.

To match the correlation between child skill ranks across school periods, the model requires large child skill shocks, especially during primary school, with a standard deviation of 0.052. The estimated own skill elasticity increases between primary school ages (0.65) and end-of-secondary school ages (0.81). At the same time, the parental education intercept in the child skill technology decreases from 0.072 to 0.032.

4.5 Validation Exercises

We assess the model's validity using two approaches. First, as is standard in the literature, we compare non-targeted moments from our model simulated data to their counterparts in the NEPS data or using estimates from other research papers. Second, we investigate the effects of school track choice on later-in-life economic outcomes for a set of *marginal* students and compare the results to the null effects reported in Dustmann et al. (2017) for Germany.

Non-targeted Moments

We summarize selected non-targeted moments and their data or external counterparts in Table 6. The first set of moments pertains to child skills. Our model features slightly smaller differences in average child skills by parental education and comparable differences in average skills by school track. In both data and model, these differences increase between primary and secondary school, before staying relatively constant.[40] In general, differences across school tracks are larger than differences across parental education. Lastly, our model produces realistic child skill rank-rank correlations *within* school tracks.

The second set of moments concerns the influence of parental education and child skills on educational choices. The data shows that the share of college-educated parents who send their child to an academic track school is around 66%, which aligns with our model prediction (70%). For non-college-educated parents, this share is only 32% (30% in the model). Moreover, we regress a dummy variable that equals one if a child attends an academic track school on the child's skills before secondary school to assess the skill gradient in academic track choice. The estimated coefficient is around 0.9 in the data and 1.0 in the model, sug-

[40] See for instance Passaretta et al. (2022); Nennstiel (2022); Schneider and Linberg (2022) who investigate the NEPS data and find stable or growing socioeconomic status gaps in children's skills.

Table 6: Non-targeted moments

Moment	Data	Model
Child Skill Moments		
Differences in average skills by parental education (in standard deviation)		
Primary School	0.53	0.44
Beginning Secondary School	0.66	0.53
Middle Secondary School	0.71	0.54
Differences in average skills by school track (in standard deviation)		
Beginning Primary School*	0.84	0.80
Beginning Secondary School	1.10	1.14
Middle Secondary School	1.11	0.95
Rank-rank coefficients		
$\text{Rank}_{j=2} - \text{Rank}_{j=3}$ if S = 1*	0.62	0.66
$\text{Rank}_{j=3} - \text{Rank}_{j=4}$ if S = 1	0.68	0.73
$\text{Rank}_{j=2} - \text{Rank}_{j=3}$ if S = 0*	0.64	0.67
$\text{Rank}_{j=3} - \text{Rank}_{j=4}$ if S = 0	0.74	0.74
Skill evolution during secondary school		
Reg. θ_4 on θ_3 and E: coef. θ_4	0.81	0.66
Reg. θ_4 on θ_3 and E: coef. E	0.04	0.04
Education Choices		
% in academic track if college parents	66%	70%
% in academic track if non-college parents	32%	30%
Reg. E' on θ_4 & S: coef. S	0.41	0.39
Reg. E' on θ_4 & S: coef. θ_4	0.395	0.495
Intergenerational Mobility and Inequality		
Parental Income Gradient (Dodin et al., 2021)	0.52	0.32
Q5/Q1 A-track on income (Dodin et al., 2021)	2.13	1.82
Q1 A-track on income (Dodin et al., 2021)	0.34	0.30
Gini Coefficient of Income	0.29	0.26
College Wage Premium	1.35	1.46

Notes: This table presents the non-targeted moments and their model-generated counterfactuals.
* We exploit the panel structure of the datasets and group students by future school track assignation.

gesting that our model generates a realistic importance of child skills for the track choice. Given that we target the college rates (shares) by parental education and the transition rates from both school tracks to college, it is also no surprise that we match the coefficient on academic track from a similar regression of college education on school track and skills, at around 0.4.

The third set of moments relates to further measures of intergenerational mobility and cross-sectional inequality. To assess the model's validity here, we compare its implications vis-à-vis the estimates on social mobility in Germany reported in Dodin et al. (2021). Using a different data set than we, they regress a dummy of academic-track school graduation of a child on the percentile income rank of her parents, finding that a ten-percentile increase in the parental rank is associated with a 5.2 percentage point increase in the probability of graduating from an academic track school. In our model, a comparable estimate yields a 3.2 percentage point increase. Moreover, Dodin et al. (2021) report absolute graduation rates for children from the first quintile of the income rank distribution (Q1) of 34% and a ratio of the fifth income rank quintile over the first quintile of 2.13. Our model-generated data squares well against these external estimates (30% and 1.82, respectively).

Long-term effects of Track Choice for Marginal Students

Dustmann et al. (2017) analyze the long-term labor market effects of early school track choice in Germany using a quasi-experimental setting. Their identification strategy makes use of the existence of a (fuzzy) cut-off age for school entry in the German system. Children born just before the cut-off age are less likely to go to an academic track secondary school simply because they are younger and, therefore, less developed than their class peers at the time of the track decision. This induces a quasi-randomness in secondary school track choice based on the date of birth. The authors then investigate the effect of that date of birth on later-in-life wages, employment, and occupation. They find no evidence that the track attended in secondary school affects these outcomes for the marginal children around the school entry cut-off.[41]

We use our model-simulated data to perform a similar exercise. In particular, we are interested in comparing the later-in-life outcomes of children who are very similar in terms of their state variables at the point of school track choice but end up going to different school tracks. To that end, we calculate the average present values of lifetime income, and average

[41] Note that Dustmann et al. (2017) control for the effect that being born after the cut-off age directly harms a child's later wages since it means that her labor market entry is later so that at any given age, she will have accumulated less work experience.

present value of lifetime wealth conditional on all states prior to entering secondary school – parental human capital, assets, education, and learning ability and child skills – of children who go to an academic track school and children who go to a vocational track school.[42] Conditional on all other states, differences in the track allocation can only arise due to the stochastic track utility shock, which we can also interpret as arising from age-at-school-entry effects.

We find that going to the academic track instead of the vocational track is associated with a 6.6% higher present value of lifetime labor income, and a 4.4% higher present value of lifetime wealth for these, otherwise very similar children. While not zero, these differences seem relatively small in relation to overall inequality in these outcomes. For example, the 6.6% higher present value of lifetime labor income is around 1/10th of a standard deviation of lifetime labor income. Moreover, in our model, the track choice is only between one vocational and one academic track, whereas Dustmann et al. (2017) consider three tracks, of which two can be classified as vocational. We would generally expect children at the margin of these two vocational tracks to show fewer differences in lifetime outcomes. In sum, we conclude that the implications our model entails with respect to the effect of tracking on *marginal* children are not at odds with the reduced-form evidence presented in Dustmann et al. (2017).

5 Quantitative Results

Our model allows us to understand the effects of school tracking not only for marginal children but for the whole distribution of children, their educational and labor market outcomes, as well as their economic mobility relative to their parents. To that end, we first quantify the main determinants of educational choices and their consequences for lifetime inequality. We then use the model to study the effects of a counterfactual policy reform that postpones the school tracking age to 14 or abolishes tracking altogether on aggregate and distributional outcomes across generations. Finally, we perform counterfactual analyses of economies in which we reduce the direct parental influence on the school track choice.

[42]Concretely, we partition all continuous states into 10 groups of equal size each. Lifetime labor income is computed as the discounted sum of all labor income during the adult periods, and lifetime wealth is that sum plus the initial monetary transfer from the parent to their independent child.

5.1 Determinants and Consequences of Educational Choices

In this section, we use our calibrated model to investigate the main determinants of educational choices and quantify their consequences for lifetime economic outcomes.

5.1.1 Determinants of the School Track Choice

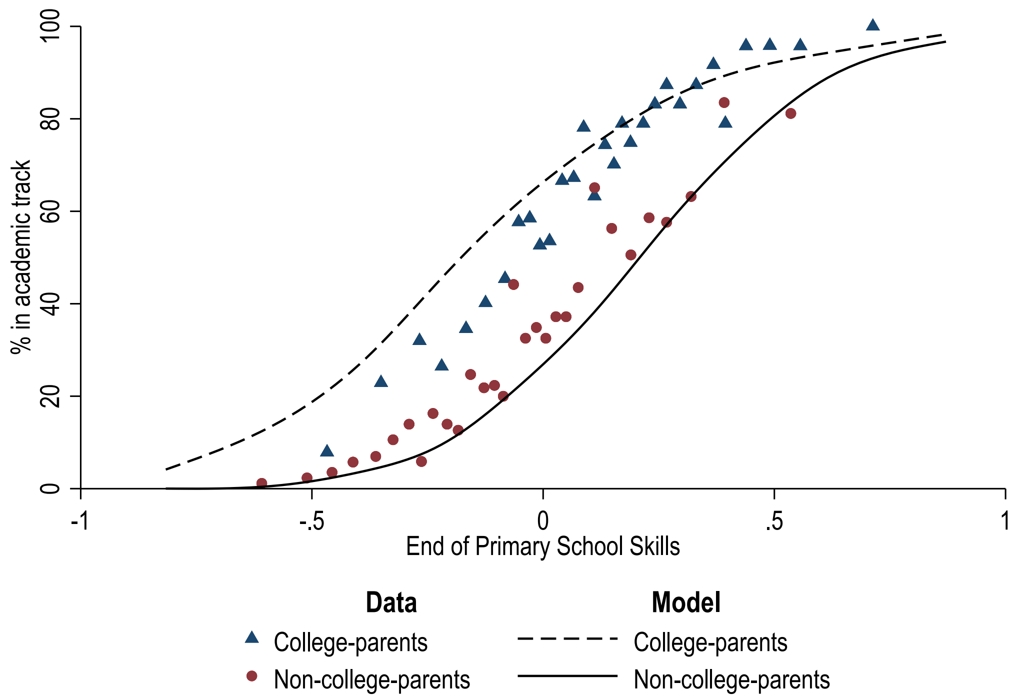
Our calibrated model predicts that children's skills largely determine the school track choice. Figure 2 shows the relationship between skills and the academic track choice, separately by parental education. The model-generated data matches remarkably well the increasing, S-shaped probability of academic-track attendance in skills observed in the NEPS data. Parental education is another important independent driver of the school track choice. Table 7 reports the coefficients from a regression of an academic school track dummy on all states at the time of the tracking decision in our model. Column (1) indicates that while child skills at the time of the track choice, θ_3 have the strongest impact on the track decision, coming from a parent with college education increases the probability of going to college by 42 percentage points, which is around half the impact of a one standard deviation increase in child skills.

In the model, parental education can influence the track choice, net of the effects coming through child skills, human capital, or wealth, in three ways. First, college-educated parents know their children learn faster than their non-college-educated counterparts. This comes from the estimated direct parental education effect in the child skill production technology, ω_5. This knowledge may prompt college parents to send their child to the academic track even if their child's skills are lower than those of a child from a vocational parent. Second, parents know their child will receive a college taste shock that depends on their parent's education, governed by μ_{ν,E^p}. In anticipation of this, college parents, for instance, may have a stronger incentive to send their child to an academic track school as this, everything else equal, increases the likelihood of college admission. Thirdly, even net of college tastes, parents face asymmetric academic track utility costs $\chi(E)$.

To understand how important each of these channels for the school track choice is, we perform a series of three counterfactual experiments (Columns (2)-(4) in Table 7), in which we isolate each effect by setting to zero the parental education effect parameter $\omega_{5,j=3,4}$, the means in college taste shocks across parental education $\Delta(\mu_{\nu,E^p})$ or the asymmetry in academic track costs χ_E.[43]

[43]In doing so, we again solve for the stationary general equilibrium, allowing prices to clear the markets and average child skills across tracks to be consistent with the parents' track decision. In Column (2) in Table

Figure 2: Probability of attending the Academic School Track



Notes: This figure shows the share of children attending the academic school track as a function of their skills. The triangles and dots are data moments and stand for children from college and non-college backgrounds, respectively; the baseline model simulated analogs are in dashed and solid lines —data source: NEPS, cohort 3. All observations are weighted so that the shares of children in each track correspond to the targeted ones.

Table 7: School Track Choice Determinants

	Dependent Variable: $S = A$			
	(1)	(2)	(3)	(4)
	Baseline	$\omega_{5,j=3,4} = 0$	$\Delta(\mu_{\nu,E^p}) = 0$	$\chi_0 = \chi_1 = 0$
θ_3	0.78	0.79	0.86	0.82
$E = 1$	0.42	0.38	0.18	0.32

Notes: This table reports the coefficient estimates of regressions of an academic school track dummy on beginning of primary school skills and parental education, controlling for all other states at the time of the tracking decision and a constant. Column (1) corresponds to the baseline economy. In Column (2), we shut down the channel of differential parental inputs in periods 3 and 4. Column (3) considers the case of identical college taste shock by parental education. In Column (4), we remove the parental preference bias for education.

In all cases, the coefficient on parental education drops, and the coefficient on child skills before the track decision increases. The magnitude of the effects, however, varies across the counterfactual scenarios. In particular, while shutting down the parental education parameter has little effect (Column (2)), shutting down the college taste shocks across parental education approximately halves the coefficient on parental education (Column (3)). Asymmetric academic school track costs also matter as shutting them down reduces the coefficient on parental education by around 24% (Column (4)).

We now turn to quantifying the consequences of educational choices for lifetime economic outcomes.

5.1.2 Sources of Inequality

In the spirit of Huggett et al. (2011) and Lee and Seshadri (2019), we can decompose how much of the variation in lifetime economic outcomes of our model agents can be explained by various factors at various ages. We focus on (the present value of) lifetime labor income and wealth as our economic outcomes of interest. We begin by computing the contribution of each state variable of a freshly independent child at age 18 to the variation in these outcomes.[44] The states are the school track in secondary school S, initial adult human capital h_5, initial transfers received from the parent a_5, the college choice E, parental college education E^p, and innate learning ability ϕ.

Table 8: Contributions to Lifetime Inequality

Life Stage	States	Share of Explained Variance: Lifetime Earnings	Lifetime Wealth
Independence (age 18)	$(S, \phi, h_5, a_5, E, E^p)$	69%	65%
	(S, ϕ, h_5, E^p)	60%	61%
	(S, ϕ, a_5, E, E^p)	49%	42%
School Track Choice (age 10)	$(S, \phi, \theta_3, h_{11}, a_{11}, E)$	30%	33%
	(S)	16%	15%
Pre-Birth (parent age 30)	(E, ϕ, h_8, a_8)	14%	21%

Notes: This table shows how much of the variation in lifetime economic outcomes is explained by different factors at different ages.

7, we isolate the effects of the first channel by solving the model with $\omega_{5,j=3,4} = 0$ yet leaving $\omega_{5,j=3,4} > 0$ in the simulation of the distribution. That is, we assume that parents do not take into account the direct effect of their own education on child skill development during secondary school when making the track decision. The skills, however, still evolve as in the baseline model.

[44] Concretely, we follow the approach in Lee and Seshadri (2019) and decompose the unconditional variances in lifetime labor income and lifetime wealth into conditional variances after conditioning on the state variables.

Row 1 of Table 8 summarizes that 69% of the variation in lifetime labor income can be accounted for by all states at the age of 18. In terms of lifetime wealth, this number is around 65%.[45] Thus, our model suggests that lifetime outcomes are already largely predetermined when agents become independent and can enter the labor market. Note that all uncertainty regarding initial human capital and the college decision has been made at this stage. The remaining unresolved uncertainty over human capital (market luck) shocks during the working years has, therefore, more minor effects on lifetime inequality. The explained share of variation in lifetime outcomes remains relatively high if we only condition on the states before the college decision has been made and the inter-vivos transfers have been realized (Row 2). This suggests that these states are not major sources of lifetime inequality. On the contrary, if we only exclude initial adult human capital h_5 (Row 3), the share of explained variance in lifetime earnings drops by 20 percentage points, and the share of explained variance in lifetime wealth by 23 percentage points. This highlights the importance of variation in initial human capital, and therefore of end-of-school skills, as a driver of lifetime inequality.

Using the same methodology, we can also evaluate how much lifetime inequality is already determined at the time of the school track choice. Conditioning on all states at that age, around a third of lifetime earnings and wealth variation is explained (Row 4). Yet the explained share is significantly smaller than after school, suggesting that the learning outcomes during secondary school play an important role in shaping later-in-life inequality. Conditioning on the initial school track choice alone can account for 16% of lifetime earnings variation and 15% of lifetime wealth variation. However, this should not be interpreted as the marginal effect of school track choice on lifetime outcomes, as the initial school track choice is, for example, highly correlated with child skills at that age.

The last row of Table 8 shows the contribution of parental states prior to the birth of their children to their children's lifetime outcomes. At this stage, uncertainty regarding child skill and human capital shocks or the child's learning ability has not yet been realized (i.e., ϕ denotes the parent's ability). Around 14% of the variance in lifetime earnings of the yet-to-be-born child is predetermined by parental education, ability, human capital, and wealth. For lifetime wealth, this share is higher at 21%, pointing to the critical role of wealth transfers. For example, using the same decomposition of the unconditional variance of transfers into parental states pre-birth, we find that more than a third of the variation in transfers (35%) is predetermined before the child's birth. In contrast, only around 23% of the variation in human capital at age 18 is predetermined before birth, highlighting the role

[45] These numbers are comparable with estimates for the U.S. (Lee and Seshadri, 2019; Huggett et al., 2011; Keane and Wolpin, 1997).

of the schooling years and shocks in shaping adult human capital.[46]

5.2 The Timing of School Tracking

In countries with an early tracking system, such as Germany, it is often argued that postponing the tracking age will improve social mobility without incurring efficiency losses (Woessmann, 2013). While some reduced-form estimates, exploiting cross-country, federal-state level, or time differences in tracking policies exist, little is known about the aggregate, distributional, and inter-generational consequences or welfare effects of a large-scale reform that changes the timing of school tracking.

To evaluate such a reform in the context of Germany, we conduct a series of counterfactual experiments using our calibrated model, in which we postpone the tracking age from ten to fourteen or abolish tracking during secondary school altogether. In each experiment, we assume that in the periods preceding tracking, all children attend a school that belongs to a comprehensive school track, just like during primary school in $j = 2$. All parameters, including those governing school academic track costs and college costs, remain the same as in the baseline economy.[47] We then compare steady-state equilibrium outcomes, which can be considered long-run outcomes of the policy change.

We present the effects of the counterfactual experiments on aggregate, distributional, and social mobility outcomes in Panel A. of Table 9. In addition, we calculate the relative changes in average welfare, defined as the percent change in consumption that a newborn in the baseline economy would require to be equally well off as in the policy counterfactual. As is common in the literature, we calculate this consumption equivalence welfare measure under the veil of ignorance, meaning that all policy functions remain unchanged.[48]

5.2.1 Postponing School Tracking by Four Years

Columns (1) and (2) present the results of postponing tracking from age ten to age fourteen, corresponding to the average tracking age in OECD countries (OECD, 2013). In Column

[46]In comparison to Lee and Seshadri (2019) in the U.S. case, our estimated contribution of parental states prior to the birth of a child to her eventual lifetime outcomes is somewhat smaller (in particular they find that almost half of the lifetime wealth variation is pre-determined at that stage. These differences may reflect that firstly, intergenerational mobility estimates in Germany tend to be smaller than in the U.S. Secondly, we incorporate explicitly the uncertainty in child skill realizations over the childhood years, while Lee and Seshadri (2019) focus on endogenous parental investments that could explain in particular the large explanatory power of pre-birth parental states for child human capital.

[47]In the case of no tracking, we assume that the fixed college utility costs $(\psi + \psi_v)$ are a weighted average of the baseline economy.

[48]Appendix Section I provides welfare definition.

Table 9: Timing of Tracking Counterfactual Experiments - Results

	Changes in %		
	(1)	(2)	(3)
Economy	PE	GE	GE
Tracking Age	14	14	Never
Panel A. - Aggregate, Distributional and Intergenerational Outcomes			
Efficiency			
Output (Y)	+0.2	-0.1	-0.2
Human Capital (H)	+0.2	-0.1	-0.4
Cross-sectional Inequality			
Gini of earnings	-0.4	-0.4	-0.8
College wage premium	-4.0	-0.2	-2.8
90th/10th percentile of income	-0.1	-0.4	-0.8
Mobility			
Intergenerational income mobility*	+3.5	+2.2	+23.9
Parental Income on Academic Track (Dodin et al., 2021)	-15	-6.8	-
Welfare	+0.18	-0.05	-0.08
Panel B. - Educational Outcomes			
% Academic track	+5.4	+2.0	
...if college parents	+2.2	+1.5	
...if non-college parents	+6.8	+2.4	
% College	+3.9	-0.3	-0.2
...if college parents	+2.9	+0.5	-18.1
...if non-college parents	+3.4	-0.9	+16.7
...if academic track	-0.4	-1.7	
...if vocational track	+3.0	+0.5	
Average end-of-school skills ($\bar{\theta}_5$)	+4.3	-1.6	-2.7
Average middle-of-school skills ($\bar{\theta}_4$)	+7.5	-2.1	-1.8
Variance of end-of-school skills ($Var(\theta_5)$)	-0.3	-0.2	-2.1
Variance of middle-of-school skills ($Var(\theta_4)$)	-0.7	-0.9	-2.7
Correlation between academic track and initial skills	-20	-14	
Correlation between end-of-school skills and initial skills	-0.5	-0.1	-3.3
Correlation between college graduation and initial skills	-18	-12.6	-69.5
Correlation between college parents and end-of-school skills	-6.0	-6.0	-26.2
Correlation between college graduation and end-of-school skills	-4.1	-3.5	-12.9

Notes: This table presents changes in outcomes in % due to delaying the school tracking choice by four years (from the age of ten to the age of fourteen) or abolishing tracking altogether. Column (1) displays percentage changes due to postponing tracking in partial equilibrium, that is, if prices are unchanged. Column (2) shows the effects of postponing tracking in general equilibrium. Column (3) presents the effect of abolishing school tracking in general equilibrium.

* Intergenerational mobility is measured as minus the income rank-rank coefficient.

(1), wages (w_0, w_1) and the interest rate r remain at the same values as in the baseline case. That is, we compare the partial equilibrium outcomes of the policy counterfactual. In Column (2), prices adjust; that is, we compare the general equilibrium outcomes of the policy counterfactual. As before, the instruction paces during all school stages are set to the level that is optimally chosen by a policymaker given the allocation of children across tracks.

We find an efficiency-equity trade-off of postponing tracking in general equilibrium but not in partial equilibrium. In partial equilibrium, Column (1), both aggregate output Y and aggregate human capital H increase by 0.2%.[49] At the same time, cross-sectional inequality, as measured by the Gini coefficient of labor income, drops by 0.4%. Similarly, the college wage premium decreases, and the ratio of the 90th to 10th percentile of income decreases. Mobility is improved, as indicated, for example, by the intergenerational elasticity of income, which drops by 3.5%. In a similar vein, the dependence of going to an academic track school on parental income drops by 15%. These effects translate into an improvement in average welfare from postponing tracking, in the range of 0.18% consumption equivalent units.

In contrast, the model-predicted effects of postponing tracking change, once we allow for the adjustment of wages on the labor market and, therefore, general equilibrium effects of the human capital changes in the economy. Column (2) of Table 9 reports that, while the gains in terms of inequality and social mobility persist, albeit at a smaller level, the effects on aggregate human capital and output actually reverse as both decrease by 0.1%. Our quantitative results therefore indicate that postponing tracking incurs a trade-off between equality and mobility improvements on the one hand, and aggregate efficiency losses on the other hand, once general equilibrium effects are taken into account.[50] Moreover, as average welfare slightly decreases relative to the early tracking benchmark, the second force seems to quantitatively outweigh the first.[51]

[49]Given that aggregate production is Cobb-Douglas in both physical capital and human capital, this implies that also aggregate physical capital increases.

[50]This result can be viewed in a similar spirit to the efficiency-mobility trade-off in Benabou (1996), who has shown that policies aimed at improving mobility may entail penalties in terms of growth, or more recently in Arenas and Hindriks (2021), who argue that more equal school opportunities by parental income raises social mobility but come at the cost of modest efficiency losses in terms of human capital.

[51]It should be noted that the standard consumption equivalent welfare measure used by us and in the related literature does not take into account improvements of intergenerational mobility that occur across cohorts. Rather the standard welfare measure (see definition in Appendix Section I) only captures the trade-off between efficiency and redistribution within cohorts. Whether our welfare conclusions regarding a postponement of the school tracking age hold also if a planner takes into account mobility is an interesting question that requires future research.

Understanding the sources of the efficiency-mobility trade-off

In our model, aggregate human capital, and hence aggregate output, is driven by the level of skills, learned during school, that translate into adult human capital and by the distribution of college education across workers. In particular, a higher share of college-educated workers has a mechanical effect on aggregate human capital in the economy through college-specific growth rates $\gamma_{j,E}$.[52] As we argued in Proposition 3 in Section 3, the effect of postponing tracking on the average skill level of children is theoretically unclear and depends in particular on the degree of uncertainty about the child skill evolution. At the same time, college parents provide high parental inputs into the child skill development, which means that a higher share mechanically boosts end-of-school skills. Ultimately, this effect is driving the efficiency gains in the partial equilibrium case. As indicated in Panel B. of Table 9 the share of college parents in steady state increases by 4.1%, and so does the share of children in academic track schools (+5.4%) and average end-of-school skills (+4.3%). In fact, because of the higher parental inputs, average child skills already before tracking at age 10 are higher in the partial equilibrium late tracking case compared to the early tracking economy. This can be seen in Figure 3, where we plot the evolution of average child skills in the early tracking economy (red), the late tracking economy in partial equilibrium (green), and the late tracking economy in general equilibrium (blue).

In the general equilibrium case, however, the college share remains approximately at its baseline level as college wages adjust downwards and non-college wages upwards. For that reason, also parental inputs are approximately the same as in the baseline, early tracking case, which results in very similar average child skills at age 10 (see Figure 3). Our calibrated model then predicts that postponing tracking leads to losses in average child skills. Concretely, average child skills in period $j = 4$, that is right before late tracking, drop by 2.1%, and end-of-school skills, drop by 1.6%, on average. As explained in Section 3, these learning losses intuitively arise from the prolonged period of comprehensive school, during which instruction becomes less efficient. Moreover, our model predicts that these losses then cannot be recuperated by learning efficiency gains that arise when more uncertainty about child skills is resolved in the late tracking case (see Figure 3).[53] Eventually, these learning losses then

[52] Labor supply can also affect aggregate human capital, but it stays approximately constant across the policy counterfactuals.

[53] In fact, the skill growth between the middle and end of secondary school in both the early and late tracking cases are quite similar. This is a consequence of the fact that the variances of skills in each track remain at a similar level compared to the early tracking case. The reason for this is an increase in the misallocation originating from parental school track preferences when tracking is postponed, which hampers learning efficiency.

translate into less human capital and explain the modest output losses that result from postponing tracking.[54]

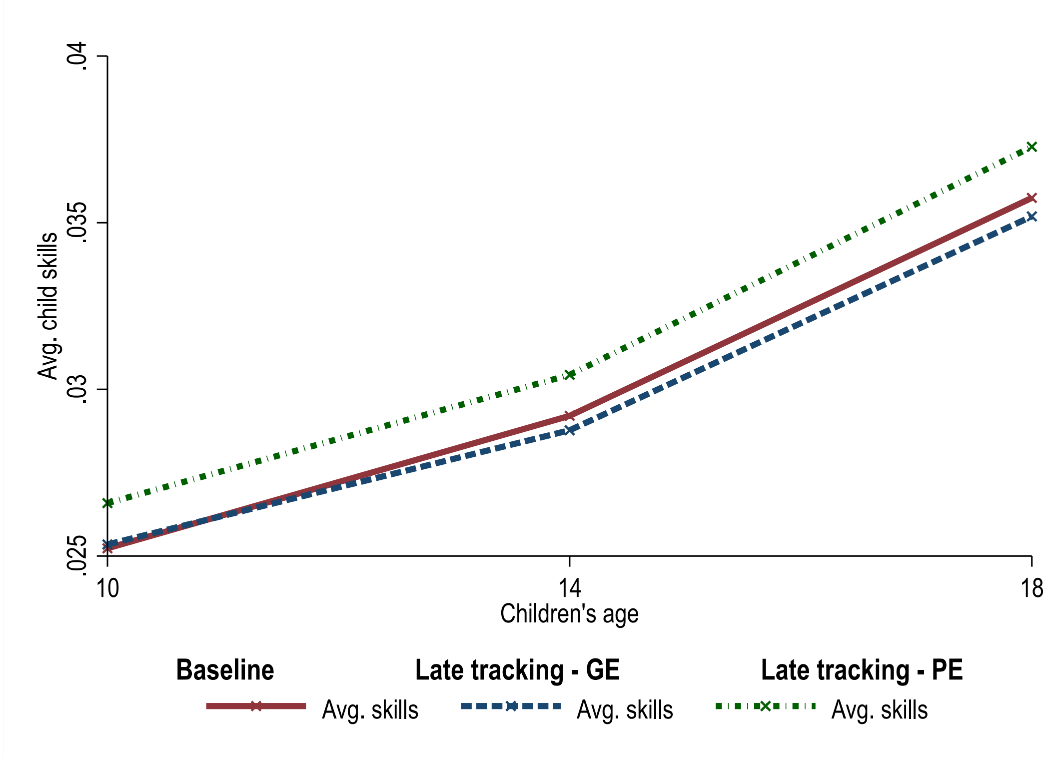
The effects of later tracking on inequality and mobility are fundamentally also rooted in the consequences of the policy change on the child skill distribution. As reported in Panel B. of Table 9, one more model period of comprehensive school decreases the overall heterogeneity in skills in the middle of secondary school (i.e. $Var(\theta_4)$ drops by 0.9%.). Intuitively, this is due to the fact that children who would have gone to a vocational track school in the early tracking economy are now exposed to, on average better peers, while children who would have gone to an academic track school are now surrounded by, on average, lower skills.[55] On top of that, one more period of comprehensive school in the late tracking case harms relatively more children from college-educated households as they would have been more likely to go to an academic track school and benefits relatively more children from non-college households, who would have been more likely to go to the vocational track and are more likely to occupy the center of the skill distribution, which as we have argued before are the children that gain most from comprehensive schooling. As a result, differences in skills between parental backgrounds decrease, and relatively more children from a non-college parental background go to an academic track school once they are tracked in the late tracking case (+2.4%) relative to children from college parents (+1.5%). This can explain the increase in mobility as measured by the dependence of academic track graduation on the parental background (Dodin et al., 2021).

The lower inequality in skills after one more period of comprehensive schooling translates into smaller differences in average skills between children in the academic and vocational track, once they are tracked. This is reinforced by the reduced differences between parental backgrounds in the track choice, and the fact that the track decision itself becomes less dependent on skills. The overall effect of postponing tracking on the child skill differences between tracks is plotted in the left panel of Figure 4, comparing the early tracking baseline economy (red) and the late tracking GE economy (blue). Smaller differences in skills across school tracks then entail smaller differences in adult human capital across college and non-college workers, which is again aided by the fact that relatively more children from a

[54]The learning losses from longer comprehensive school also serve as an explanation for the raised incentive to send a child to an academic track school, where average peer skill levels are higher, which partially compensates for less efficient learning. Similarly, this explains why, in the partial equilibrium case, the share of college workers increases in the first place. When average skill levels after school are smaller, newly independent adults can compensate for this by choosing to go to college more often.

[55]As we have shown in Proposition 2 in Section 3, the effect of tracking on the overall variance of skills depends crucially on the presence of these direct peer effects.

Figure 3: Evolution of Average Child Skills in Counterfactual Experiments



Notes: This figure shows the evolution of average child skills from age 10 to 18. The baseline model simulated data is a red-solid line; the late tracking economy in partial equilibrium in a green-dotted-dashed line; and the late tracking economy in general equilibrium in a blue-dashed line.

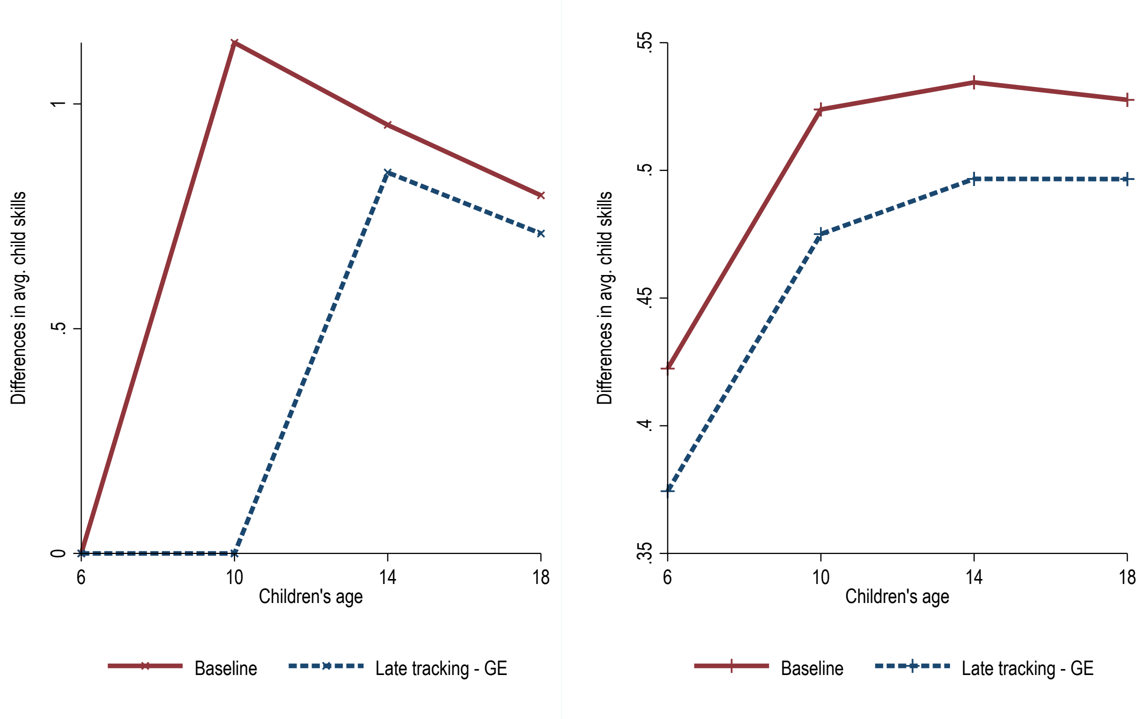
vocational school track go to college after the policy change. Given that college education and end-of-school skills and, thereby, human capital are the main determinants of income, overall income inequality therefore declines.

A consequence of these effects is that not only the school track, but also the end-of-school skills, and the probability of going to college become less dependent on the initial skills, which are transferred from parents to children (see bottom rows of Panel B. in Table 9). For that reason, intergenerational mobility, if defined as the dependence of economic outcomes of the child on parental economic outcomes, decreases. Interestingly, our model predicts that, in steady state, college attainment of children from non-college parents is still as likely or even slightly more likely than in the baseline, early tracking case. However, since college- and non-college parents become more similar in terms of their human capital, mobility in terms of income still improves.

Finally and importantly, the effects on inequality and mobility (and also efficiency) are reinforced through intergenerational linkages. For example, as college and non-college-educated

Figure 4: Differences in Average Child Skills in Early and Late Tracking Economy

(a) by school track (b) by parental education



Notes: These figures show the standardized differences in average skills between school tracks (left panel) and parental education (right panel) from age 6 to 18. The baseline model simulated data is a red-solid line, and the late tracking economy in general equilibrium is in a blue-dashed line.

parents become more similar in terms of their skills, so do their children who inherit these skills. The differences in skills in terms of standard deviations between parental backgrounds are shown in the right panel of Figure 4. As can be seen, the relative differences between children of different parental backgrounds in the late-tracking GE case are reduced already at age 6.

Comparison of Effects on Learning Outcomes to Empirical Estimates

Generally, empirical estimates of the effects of between-school tracking policies on the average learning outcomes of children offer no clear consensus, as identification of causal effects is made difficult by severe endogeneity issues (Hanushek and Wössmann, 2006). Using the same dataset as we do, Matthewes (2021) compares the mathematics and reading achievement outcomes of school children in non-academic school tracks in Germany who benefit from two more years of comprehensive school in some federal states versus those that are already tracked in these years in a difference-in-differences framework. He finds that later tracking even improves average achievement outcomes. However, in contrast to our setup, his analysis does not consider children in academic track school who are already separated but compares children in two different non-academic tracks.[56] The results in Matthewes (2021), therefore, are not directly informative about the effects of a broad comprehensive school reform, which places *all* school children in schools of the same track for a longer period of time. However, they suggest that the effects of tracking may be heterogeneous in the sense that it could be particularly children from lower socio-economic backgrounds that benefit from de-tracking reforms.

This is corroborated by empirical evidence from Scandinavian countries, who have all undergone comprehensive school reforms in the last 60 years and often find that longer comprehensive schooling decreases the effect of family background on educational attainment (see, for instance Meghir and Palme, 2005; Aakvik et al., 2010; Pekkala Kerr et al., 2013).[57] Similarly to Meghir and Palme (2005), who study an increase of compulsory schooling to nine years from seven or eight years in Sweden, we find a negative effect of our policy on attainments for children from college parents (-3% in end-of-school skills) but a positive effect for children from non-college parents (+5% in end-of-school skills). Our results are also in line with evidence that attendance at academic track schools becomes less dependent on the parental background when tracking occurs later. In particular, we find an increase in academic shares for college and non-college parents' children, but relatively more so for non-college parents' children.

In terms of the effects of school tracking policies on inequality in learning outcomes, most existing evidence comes from comparisons of early and late tracking systems across countries (Hanushek and Wössmann, 2006; Brunello and Checchi, 2007).[58] They consistently find that tracking raises educational inequality as measured by child achievement test scores. Our result of lower heterogeneity in child skills during secondary school is thus in line with these findings.

Finally, while the reduced impact of the family background on educational attainment in

[56] Traditionally, the school system in many federal states in Germany consisted of three tracks. One academic track (*Gymnasium*) and two non-academic tracks that differ much less in terms of their curriculum than between academic and non-academic tracks.

[57] Since the reforms in these countries came together with other education policy changes, in particular with more mandatory schooling years, the estimated effects can often not be unequivocally attributed to the tracking regime change. Studies about de-tracking reforms in Britain (e.g. Pischke and Manning, 2006) also often arrive at mixed results. Piopiunik (2014), who study an *increase* in tracking in one of the federal states in Germany, Bavaria, also led to learning losses for the lower-skilled children.

[58] As pointed out by Waldinger (2006) or Betts (2011), these studies often come with significant identification challenges given the unclear classification of countries in early and late tracking systems or the possibility of unobserved differences driving the results.

secondary school is often already interpreted as evidence for improvements in social mobility following de-tracking reforms, it does not necessarily follow that such improvements lead to lower association between child and parental outcomes later in life.[59] Similarly, smaller inequality in test scores does not necessarily need to translate into lower cross-sectional inequality in terms of labor market outcomes. As argued before, an assessment of the effects of school tracking policies on these outcomes is challenging as it requires the consideration of general equilibrium effects that a change in the skill composition of students may entail on the labor market. Most existing empirical evidence, however, comes from relatively short-term evaluations of tracking reforms that cannot consider such effects.

To the best of our knowledge, the only empirical estimates on the effect of a broad comprehensive school reform on the intergenerational elasticity of income come again from the Nordic countries (Holmlund, 2008; Pekkarinen et al., 2009). In particular, the reform in Finland, undertaken subsequently across regions in the 1970s, is similar in scope to our experiment as it postponed the track allocation from age 11 until age 16. Pekkarinen et al. (2009) find that the elasticity between fathers' and sons' relative earnings declined by seven percentage points due to the reform (from 0.3 to 0.23). Our model also predicts a decrease in the intergenerational income elasticity, yet the effect is quantitatively smaller. Some of this difference is likely due to the fact the reform in Finland simultaneously also changed the average length of schooling, and the content of the curriculum in schools towards a more academic orientation and went from a largely private to a public school system.

In sum, we maintain that our model's predictions of the effect of a postponement of the school tracking age by four years in Germany are not at odds with existing empirical evidence. In particular, our model predicts relative learning gains for children from lower socio-economic backgrounds and a decreased dependence on educational outcomes in secondary school on family background, which are among the most robust empirical findings in the literature. The fact that our model predicts average learning losses from such a pervasive school tracking age reform can, in our eyes, not be refuted by existing evidence (nor can it be corroborated). It ultimately rests on the assumption of complementarity between child skills and the teaching practices in school, as highlighted in Section 3, which is itself based

[59]For example, Malamud and Pop-Eleches (2011) analyses the effects of school-tracking age postponement in Romania. While they find that children from disadvantaged backgrounds were significantly more likely to attend and graduate from academic track schools following the reform, this did not lead to an increased share in the college graduation probabilities of these disadvantaged children, which they attribute to the same overall share of college slots available pre- and post-reform. The quantitative results of our model similarly predict, that while postponing tracking increases mobility in school track choice, this does not lead to higher mobility in college attainment. This effect is driven by parental-education-specific college tastes.

on empirical evidence (e.g. Duflo et al., 2011; Aucejo et al., 2022). The strength of our model-based approach is that it informs not just the short-term effect of school tracking on learning and educational outcomes of school children, but how these translate into higher education and labor market outcomes within and across generations.

5.2.2 Abolishing School Tracking

Column (3) of Table 9 reports the results of a counterfactual economy, in which we abolish tracking altogether while allowing wages and the interest rate to adjust. All children go to comprehensive schools for the entirety of their schooling years, and instruction occurs at the same pace that is optimal for the overall average skill level. As a consequence, child skills become significantly more equal (i.e. $Var(\theta_5)$ drops by 2.1%). Moreover, as parents can no longer influence their children's skill evolution by choosing a specific school track, the correlation between parental background and end-of-school skills drops sharply (-26.2%). As a result, and despite college-specific preferences, mobility in higher education also increases; children from non-college parents are 16.7% more likely to graduate from college than in the baseline economy, and children from college parents are 18.1% less likely to do so. Overall, mobility as measured by the (negative of the) intergenerational income elasticity improves significantly (+23.9%).

Similarly, a completely comprehensive school system reduces cross-sectional inequality markedly. Concretely, the Gini coefficient of earnings drops by 0.8%, as does the ratio of the 90th to 10th percentile of income. In addition, the differences in human capital across college and non-college workers become smaller, which decreases the college wage premium by 2.8%.

On the other hand, abolishing tracking altogether makes learning even less efficient relative to the late tracking economy. In particular, end-of-school skills are around -2.3% smaller in this economy, on average. Eventually, this leads to losses in aggregate human capital (-0.4%) and output (-0.2%). Thus, despite considerable mobility and equality gains, a completely comprehensive school system worsens average welfare by 0.08%.

5.3 Limiting Parental Influence in the School Track Choice

In this section, we evaluate the effects of reducing parental influence on the school track choice without modifying the timing of school tracking. As discussed in Section 3, any force that impacts the school track allocation net of child skills can, in theory, be detrimental to the efficiency of skill development in secondary school if they dilute the homogeneity of peer groups in each track.[60] However, an important question is whether the consequences of such "misallocation" effects are visible not only in terms of child skill outcomes but also in the aggregate and distributional outcomes in the economy. Our model provides a suitable environment to investigate such effects.

We evaluate two counterfactual scenarios: first, we shut down the asymmetry in academic track utility costs faced by parents of different education levels ($\chi_0 = \chi_1 = 0$). As argued before, this asymmetry is a parsimonious way of capturing multiple reasons why parents systematically bias the school track choice toward their own educational path.[61] Second, we enforce that the school track allocation is governed exclusively by a sharp skill threshold, such that all children with skills below the threshold are allocated to the vocational track, while all children with skills above the threshold go to the academic track, regardless of the parental background. This threshold is chosen, such that the overall share of children in the academic track is constant relative to the baseline economy.[62]

Table 10 shows that shutting down the asymmetry in academic track utility costs or enforcing a tracking threshold improves aggregate output. As before, this is mirrored by an increase in aggregate human capital in the economy, in both cases. Note that both the share of college-educated agents and the share of children in academic track schools remains constant relative to the baseline case (see Panel B. of Table 10). The reason for the positive effects on human capital and output becomes clear when we study the distribution of skills

[60] Suppose for example, college-educated parents send their children to an academic track school, despite the fact that their skill level would optimally suggest the vocational track. In that case, this will not only harm their child's development but also cause the instruction pace in that track to adjust. This, in turn, harms the average learning gains of everyone in that track. The same effect occurs in the vocational track school if parents from non-college backgrounds send their overqualified children there purely based on preferences.

[61] We focus on this experiment as we view this as being the easiest to address by policies. For example, if the asymmetric school track costs are coming from information frictions, mentoring programs have proven very effective and almost cost-free in alleviating some of these frictions, as argued by Resnjanskij et al. (2021). While in the counterfactual scenario, we diminish parental influence from all socioeconomic groups, mentoring programs mostly target at-risk youths. Interventions that target all socioeconomic groups are rarer. An exception is Hakimov et al. (2022) who provide information about their chances of success at graduating college to all children, independently of their socioeconomic backgrounds. They find a reduction in the social elite college admission gap mostly driven by an increase in the admission of high-achieving low SES students to elite colleges.

[62] As derived in Section 3, the optimal tracking policy from the point of view of a policymaker who is only interested in maximizing aggregate end-of-school skills and cannot condition on the parental background, would be to track children at a threshold that is exactly equal to the average child skill level prior to the track decision. Given that the distribution of child skills is quite symmetric, this would result in a roughly equal split of children between tracks, which ensures that the variance of child skills in each track is minimized. However, to be comparable to the baseline economy, we select a threshold that will result in the top 44% of children in terms of their skills being allocated to the academic track and the rest to the vocational track.

Table 10: Effects of School Track Choice Counterfactuals

	Changes in %	
	(1)	(2)
	$\chi_0 = 0$ $\chi_1 = 0$	Skill Threshold
Panel A. - Aggregate, Distributional and Intergenerational Outcomes		
Output (Y)	+0.04	+0.12
Human Capital (H)	+0.05	+0.15
Gini of earnings	0.0	+0.8
Intergenerational income mobility*	+0.9	-6.5
Parental Income on Academic Track (Dodin et al., 2021)	-25	+34
Welfare	+0.04	-0.01
Panel B. - Educational Outcomes		
% Acadmic track	-0.7	0.0
...if college parents	-8.6	-12
...if non-college parents	+9.6	+15.6
% College	0.3	0.0
...if college parents	-3.5	-8.8
...if non-college parents	+ 3.9	+7.2
...if academic track	+0.5	-3.5
...if vocational track	0.0	+14.4
Average end-of-school skills ($\bar{\theta}_5$)	+0.8	+3.0
Average middle-of-school skills ($\bar{\theta}_4$)	+0.2	+4.9
Average skills in V-Track upon tracking ($\bar{\theta}_3 \vert S = V$)	-0.4	-50.0
Average skills in A-Track upon tracking ($\bar{\theta}_3 \vert S = A$)	+1.2	+38.5
Variance of end-of-school skills ($Var(\theta_5)$)	+0.2	+1.9
Variance of middle-of-school skills ($Var(\theta_4)$)	-0.2	+0.9
Variance in V-Track upon tracking ($Var(\theta_3 \vert S = V)$)	-0.4	-39.1
Variance in A-Track upon tracking ($Var(\theta_3 \vert S = A)$)	-0.4	-18.1
Correlation between A-Track and Skills in period 3	+5.7	+59.4
Correlation between academic track and initial skills	+5.4	+79.5
Correlation between end-of-school skills and initial skills	-0.6	+0.9
Correlation between college graduation and initial skills	+0.9	+54.1

Notes: Column (1) displays percentage changes relative to the baseline economy entailed by the absence of parental preference for education, and Column (2) displays percentage changes entailed by skill treshold-rule for school tracking. All results are coming from the new general equilibrium distribution.
* Intergenerational mobility is measured as minus the income rank-rank coefficient.

in the counterfactual experiments. In particular, both counterfactual scenarios lead to an increase in average child skills at the end of secondary school. This increase arises from the fact that the variation in child skills within the school tracks becomes smaller, both when eliminating parental track preferences and especially when enforcing a tracking threshold. Since lower heterogeneity in child skills within a school track improves learning efficiency, as derived in Section 3, this leads to higher average skills at the end of school. This is consistent with the explanation of the learning efficiency-reducing misallocation effects that arise when parental background or any other factors drive the school track choice independently from skills. Unsurprisingly, without asymmetry in parental academic track costs and even more so with a sharp, purely skill-based allocation rule, the correlation of school track with parental education decreases, and skills themselves become more important in explaining the track choice.

However, while mobility increases in the first counterfactual experiment (Column (1)), it increases when introducing a strict skill threshold (Column (2)). One reason for this is that an economy with a strict skill-based separation of children makes the school track, end-of-school skills and college outcomes significantly more dependent on the initial skill level, inherited from the parents (see the last three rows in Panel B. of Table 10). On top of that, a cut-off based school track allocation benefits predominantly those children in academic track schools. The argument is similar to Proposition 2 in Section 3: When factors other than skills determine the track choice, child skills in each track become more heterogeneous. In some sense, each track is thus more like a comprehensive school. As argued in Proposition 2, the learning losses from moving towards a stricter tracking system relative to a more comprehensive system are asymmetric and concentrated in the lower track, whenever the direct peer effects are positive, which is the case. Quantitatively, this effect can be seen in Panel B. in Table 10, where average skills in the vocational track at the point of the track decision actually decrease while they increase in the academic track.

Moreover, this also translates into higher earnings inequality in the counterfactual with a skill threshold compared to the baseline economy, which rises by (0.8%). As a consequence, and perhaps surprisingly, overall welfare in this counterfactual economy is slightly lower than in the baseline. In contrast, income inequality does not increase when shutting down the parental track preferences directly, even though end-of-school skills are slightly more dispersed. Given that aggregate human capital and output are higher, welfare is also increased by around 0.04% in this scenario.

In sum, these results point to an important role that measures such as mentoring programs, which have been shown to alleviate the causal influence of family background on school track decisions that is not justified by skill selection, can play in improving both efficiency and mobility at the same time. In contrast, reverting back to a purely merit-based school track selection is, according to the predictions of our model, not welfare enhancing as parents cannot choose the school track anymore, and it dampens equality and social mobility.

6 Conclusion

How important is the design of education policies for the macroeconomic analysis of inequality and social mobility? This paper argues that school tracking, a common policy across advanced countries, influences not only equality of educational opportunities for children from different parental backgrounds, but also shapes aggregate human capital and, consequently, aggregate economic efficiency. We add a macroeconomic perspective to the literature by building a macroeconomic GE model of overlapping generations that specifically zooms in on the children's schooling years. The key ingredient in the model is a parsimonious theory of child skill formation during the school years, where child skills depend on her classroom peers and on the instruction pace specific to each school track. This entails theoretical implications for the effect of school tracking policies on the distribution of child skills that align with the most robust empirical findings of the effect of tracking on educational achievements, as well as the most popular arguments in the public debate about tracking.

We tailor the model to fit the German Education System, where the track decision occurs at the age of ten of the child, and calibrate it using German data. Our calibrated model predicts that the timing of the school tracking age entails a macroeconomic trade-off between efficiency and social mobility. Concretely, a policy reform that delays the school tracking decision by four years (to age 14) in Germany leads to aggregate output losses, in the long run, that amount to around 0.1% of GDP while decreasing the inter-generational income elasticity by around 2.2%, thereby improving social mobility. Key in the evaluation of this trade-off is the consideration of general equilibrium effects in the labor market that affect the incentives governing the school track choice. The output losses from this reform fundamentally stem from learning efficiency losses due to the prolonged time of comprehensive schooling. At the same time, the social mobility gains result as longer comprehensive schooling reduces the differences in child skills across tracks and parental background, such that the eventual human capital of children depends less on their initial conditions.

We also find that reducing the direct influence of parental background on the school track leads to improvements in both social mobility and economic output, improving aggregate welfare by 0.04%. This last result points to the important role that measures such as mentoring programs, which alleviate the influence of family background on school track decisions that are not justified by skill selection (see, for instance, Raposa et al. (2019) for a review, and Resnjanskij et al. (2021) for an example in Germany), can play by simultaneously improving efficiency and mobility.

References

Aakvik, A., K. G. Salvanes, and K. Vaage (2010): "Measuring heterogeneity in the returns to education using an education reform," *European Economic Review*, 54, 483–500.

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2019): "Education policy and intergenerational transfers in equilibrium," *Journal of Political Economy*, 127, 2569–2624.

Agostinelli, F. (2018): "Investing in children's skills: An equilibrium analysis of social interactions and parental investments," *Unpublished Manuscript, University of Pennsylvania.*

Agostinelli, F., M. Doepke, G. Sorrenti, and F. Zilibotti (2023): "It takes a village: the economics of parenting with neighborhood and peer effects," Working Paper w27050, National Bureau of Economic Research.

Agostinelli, F., M. Saharkhiz, and M. Wiswall (2019): "Home and School in the Development of Children," Working Paper w26037, National Bureau of Economic Research.

Agostinelli, F. and M. Wiswall (2016): "Estimating the technology of children's skill formation," Working Paper w22442, National Bureau of Economic Research.

Arenas, A. and J. Hindriks (2021): "Intergenerational mobility and unequal school opportunity," *The Economic Journal*, 131, 1027–1050.

Aucejo, E., P. Coate, J. C. Fruehwirth, S. Kelly, and Z. Mozenter (2022): "Teacher effectiveness and classroom composition: Understanding match effects in the classroom," *The Economic Journal*, 132, 3047–3064.

Ballou, D. (2009): "Test scaling and value-added measurement," *Education finance and Policy*, 4, 351–383.

Bauer, P. and R. T. Riphahn (2006): "Timing of school tracking as a determinant of intergenerational transmission of education," *Economics Letters*, 91, 90–97.

Becker, G. S. and N. Tomes (1979): "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Political Economy*, 87, 1153–1189.

——— (1986): "Human capital and the rise and fall of families," *Journal of labor economics*, 4, S1–S39.

Bellenberg, G. and M. Forell (2012): "Schulformwechsel in Deutschland," *Durchlässigkeit und Selektion in den 16 Schulsystemen der Bundesländer innerhalb der Sekundarstufe I.*

Benabou, R. (1996): "Equity and efficiency in human capital investment: the local connection," *The Review of Economic Studies*, 63, 237–264.

Betts, J. R. (2011): "The economics of tracking in education," in *Handbook of the Economics of Education*, Elsevier, vol. 3, 341–381.

Blossfeld, H., H. Rossbach, and J. von Maurice (2019): "Education as a lifelong process: The German National Educational Panel Study (NEPS)," *Edition ZfE*.

Bonesrønning, H., H. Finseraas, I. Hardoy, J. M. V. Iversen, O. H. Nyhus, V. Opheim, K. V. Salvanes, A. M. J. Sandsør, and P. Schøne (2022): "Small-group instruction to improve student performance in mathematics in early grades: Results from a randomized field experiment," *Journal of Public Economics*, 216, 104765.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008): "The economics and psychology of personality traits," *Journal of human Resources*, 43, 972–1059.

Brunello, G. and D. Checchi (2007): "Does school tracking affect equality of opportunity? New international evidence," *Economic policy*, 22, 782–861.

Brunello, G., M. Giannini, and K. Ariga (2007): "The optimal timing of school tracking: a general model with calibration for Germany," *Schools and the equal opportunity problem*, 129–156.

Brunello, G., L. Rocco, K. Ariga, and R. Iwahashi (2012): "On the efficiency costs of de-tracking secondary schools in Europe," *Education Economics*, 20, 117–138.

Capelle, D. (2022): "The Great Gatsby goes to College: Tuition, Inequality and Intergenerational Mobility in the U.S." Working paper.

Carlana, M., E. La Ferrara, and P. Pinotti (2022): "Goals and gaps: Educational careers of immigrant children," *Econometrica*, 90, 1–29.

Caucutt, E. M. and L. Lochner (2020): "Early and late human capital investments, borrowing constraints, and the family," *Journal of Political Economy*, 128, 1065–1147.

Ciccone, A. and G. Peri (2005): "Long-run substitutability between more and less educated workers: evidence from US states, 1950–1990," *Review of Economics and Statistics*, 87, 652–663.

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

Cunha, F., J. Heckman, and S. M. Schennach (2010): "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, 78, 883–931.

Daruich, D. (2022): "The Macroeconomic Consequences of Early Childhood Development Policies," Working Paper 2018-29, FRB St. Louis.

Dodin, M., S. Findeisen, L. Henkel, D. Sachs, and P. Schüle (2021): "Social Mobility in Germany," Discussion Paper DP16355, CEPR.

Doepke, M. and F. Zilibotti (2017): "Parenting with style: Altruism and paternalism in intergenerational preference transmission," *Econometrica*, 85, 1331–1371.

Dohmen, D., M. Thomsen, G. Yelubayeva, and R. Ramirez (2019): "Ermittlung der Lebenshaltungskosten von Studierenden: Aktualisierte Berechnung anhand der 21. Sozialerhebung des Deutschen Studentenwerks," *FiBS - Forschungsinstitut für Bildungs- und Sozialökonomie*.

Duflo, E., P. Dupas, and M. Kremer (2011): "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya," *American Economic Review*, 101, 1739–74.

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

Dustmann, C., P. A. Puhani, and U. Schönberg (2017): "The long-term effects of early track choice," *The Economic Journal*, 127, 1348–1380.

Epple, D., E. Newlon, and R. Romano (2002): "Ability tracking, school competition, and the distribution of educational benefits," *Journal of Public Economics*, 83, 1–48.

Epple, D. and R. Romano (2011): "Peer effects in education: A survey of the theory and evidence," in *Handbook of Social Economics*, Elsevier, vol. 1, 1053–1163.

Esser, H. and J. Seuring (2020): "Kognitive Homogenisierung, schulische Leistungen und soziale Bildungsungleichheit," *Zeitschrift für Soziologie*, 49, 277–301.

Fuchs-Schündeln, N., D. Krueger, A. Kurmann, E. Lale, A. Ludwig, and I. Popova (2023): "The fiscal and welfare effects of policy responses to the covid-19 school closures," *IMF Economic Review*, 1–64.

Fuchs-Schündeln, N., D. Krueger, A. Ludwig, and I. Popova (2022): "The long-term distributional and welfare effects of Covid-19 school closures," *The Economic Journal*, 132, 1647–1683.

Fujimoto, J., D. Lagakos, and M. Vanvuren (2023): "Aggregate and Distributional Effects of 'Free' Secondary Schooling in the Developing World," Working Paper w31029, National Bureau of Economic Research.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019): "The German socio-economic panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 239, 345–360.

Hakimov, R., R. Schmacker, and C. Terrier (2022): "Confidence and college applications: Evidence from a randomized intervention," Working paper, WZB Discussion Paper.

Hanushek, E. A. and L. Wössmann (2006): "Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries," *The Economic Journal*, 116, C63–C76.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal tax progressivity: An analytical framework," *The Quarterly Journal of Economics*, 132, 1693–1754.

Heckman, J. and S. Mosso (2014): "The economics of human development and social mobility," *Annu. Rev. Econ.*, 6, 689–733.

Henninges, M., C. Traini, and C. Kleinert (2019): "Tracking and Sorting in the German Educational System," Working Paper 83, Leibniz Institute for Educational Trajectories (LIfBi).

Holmlund, H. (2008): "Intergenerational Mobility and Assortative Mating: Effects of an Educational Reform. CEE DP 91." *Centre for the Economics of Education (NJ1).*

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of lifetime inequality," *American Economic Review*, 101, 2923–54.

Jang, Y. and M. Yum (2022): "Aggregate and Intergenerational Implications of School Closures: A Quantitative Assessment," Working Paper 234v1, CRC TR 224.

Keane, M. P. and K. I. Wolpin (1997): "The career decisions of young men," *Journal of Political Economy*, 105, 473–522.

Kindermann, F., L. Mayr, and D. Sachs (2020): "Inheritance taxation and wealth effects on the labor supply of heirs," *Journal of Public Economics*, 191, 104127.

Kotera, T. and A. Seshadri (2017): "Educational policy and intergenerational mobility," *Review of economic dynamics*, 25, 187–207.

Krueger, D. and A. Ludwig (2016): "On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium," *Journal of Monetary Economics*, 77, 72–98.

Kyzyma, I. and O. Groh-Samberg (2018): "Intergenerational Economic Mobility in Germany: Levels und Trends," Working paper, DIW.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2018): "Life cycle wage growth across countries," *Journal of Political Economy*, 126, 797–849.

Lavy, V., M. D. Paserman, and A. Schlosser (2012): "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom," *The Economic Journal*, 122, 208–237.

Lee, S. Y. and A. Seshadri (2019): "On the intergenerational transmission of economic status," *Journal of Political Economy*, 127, 855–921.

Mahler, L. and M. Yum (2023): "Lifestyle Behaviors and Wealth-Health Gaps in Germany," Tech. rep., Available at SSRN 4034661.

Malamud, O. and C. Pop-Eleches (2011): "School tracking and access to higher education among disadvantaged groups," *Journal of Public Economics*, 95, 1538–1549.

Masters, G. N. (1982): "A Rasch model for partial credit scoring," *Psychometrika*, 47, 149–174.

Matthewes, S. H. (2021): "Better together? Heterogeneous effects of tracking on student achievement," *The Economic Journal*, 131, 1269–1307.

Meghir, C. and M. Palme (2005): "Educational reform, ability, and family background," *American Economic Review*, 95, 414–424.

Nennstiel, R. (2022): "No Matthew effects and stable SES gaps in math and language achievement growth throughout schooling: Evidence from Germany," *European sociological review*.

NEPS Network (2022): "National Educational Panel Study, Scientific Use File of Starting Cohort Kindergarten," .

Neumann, I., C. Duchhardt, M. Grüssing, A. Heinze, E. Knopp, and T. Ehmke (2013): "Modeling and assessing mathematical competence over the lifespan," *Journal for educational research online*, 5, 80–109.

OECD (2013): "PISA 2012 results: What makes schools successful (volume IV): Resources, policies and practices, PISA," Report TD/TNC 114.1481, OECD.

——— (2020): "Education Policy Outlook in Germany," *OECD Education Policy Perspectives*.

Passaretta, G., J. Skopek, and T. van Huizen (2022): "Is social inequality in school-age achievement generated before or during schooling? A European perspective," *European Sociological Review*, 38, 849–865.

Pekkala Kerr, S., T. Pekkarinen, and R. Uusitalo (2013): "School tracking and development of cognitive skills," *Journal of Labor Economics*, 31, 577–602.

Pekkarinen, T., R. Uusitalo, and S. Kerr (2009): "School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform," *Journal of Public Economics*, 93, 965–973.

Piopiunik, M. (2014): "The effects of early tracking on student performance: Evidence from a school reform in Bavaria," *Economics of Education Review*, 42, 12–33.

Pischke, J.-S. and A. Manning (2006): "Comprehensive versus selective schooling in England in Wales: What do we know?" Working Paper w12176, National Bureau of Economic Research.

Pohl, S. and C. H. Carstensen (2012): "NEPS technical report-Scaling the data of the competence tests," Working Paper 14, NEPS.

——— (2013): "Scaling of competence tests in the National Educational Panel Study-Many questions, some answers, and further challenges," *Journal for Educational Research Online*, 5, 189–216.

Raposa, E. B., J. Rhodes, G. J. J. Stams, N. Card, S. Burton, S. Schwartz, L. A. Y. Sykes, S. Kanchewa, J. Kupersmidt, and S. Hussain (2019): "The effects of youth mentoring programs: A meta-analysis of outcome studies," *Journal of youth and adolescence*, 48, 423–443.

Rasch, G. (1960): *Probabilistic models for some intelligence and attainment tests.*, ERIC.

Resnjanskij, S., J. Ruhose, S. Wiederhold, and L. Woessmann (2021): "Can Mentoring Alleviate Family Disadvantage in Adolscence? A Field Experiment to Improve Labor-Market Prospects," Tech. Rep. 277.

Restuccia, D. and C. Urrutia (2004): "Intergenerational persistence of earnings: The role of early and college education," *American Economic Review*, 94, 1354–1378.

Ruhose, J. and G. Schwerdt (2016): "Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries," *Economics of Education Review*, 52, 134–154.

Sacerdote, B. (2011): "Peer effects in education: How might they work, how big are they and how much do we know thus far?" in *Handbook of the Economics of Education*, Elsevier, vol. 3, 249–277.

Schneider, T. and T. Linberg (2022): "Development of socio-economic gaps in children's language skills in Germany," *Longitudinal and Life Course Studies*, 13, 87–120.

Waldinger, F. (2006): "Does tracking affect the importance of family background on students' test scores?" *Unpublished manuscript, LSE.*

Woessmann, L. (2013): "Die entscheidende Säule," *Wirtschaftswoche Global*, 2, 110–111.

——— (2016): "The importance of school systems: Evidence from international differences in student achievement," *Journal of Economic Perspectives*, 30, 3–32.

——— (2020): "Gleiche Chancen? Je früher, desto besser! Bildungsgerechtigkeit im deutschen Schulsystem," *lautstark*, 07, 20–22.

Yum, M. (2023): "Parental time investment and intergenerational mobility," *International Economic Review*, 64, 187–223.

A Proof of Propositions

Proposition 1

For the proof of this proposition, we denote by θ_3 the child skills at the beginning of secondary school and by θ_4 the skills at the end of secondary school. Moreover, we have assumed $\kappa = 1$, $\zeta = 0$, and $\chi = 0$ and that skills at the beginning of secondary school are normally distributed with mean zero and variance $\sigma^2_{\theta_3}$. First, we show that maximizing the aggregate end-of-school skills in a tracking system implies a threshold skill level $\tilde{\theta}_3$, such that all $\theta_3 < \tilde{\theta}_3$ go to one track, call it $S = V$ and all $\theta_1 > \tilde{\theta}_3$ go to the other track, $S = A$ (and those with $\theta_3 = \tilde{\theta}_3$ are indifferent). That is, the existence of a skill threshold is a necessary condition for optimal end-of-school skills. We restrict ourselves to the case with different instruction paces across school tracks.

To that end, it is useful to rewrite θ_4 in (2) of a child in a given school track S with instruction pace P^S using Lemma 1 as:

$$\theta_4 = \theta_3 + \alpha\bar{\theta}_3^S + \frac{\beta^2}{2\delta} + \frac{\beta\gamma\theta_3}{\delta} + \frac{\gamma^2\theta_3\bar{\theta}_3^S}{\delta} - \frac{\gamma^2(\bar{\theta}_3^S)^2}{2\delta} + \eta_4. \tag{A.1}$$

After adding and subtracting $\frac{\gamma^2}{2\delta}\theta_3^2$, this can be expressed as

$$\begin{aligned}\theta_4 &= \theta_3 + \alpha\bar{\theta}_3^S + \frac{\beta^2}{2\delta} + \frac{\beta\gamma\theta_3}{\delta} + \frac{\gamma^2\theta_3^2}{2\delta} + \eta_4 - \frac{\gamma^2}{2\delta}\left(\theta_3^2 - 2\theta_3\bar{\theta}_3^S + (\bar{\theta}_3^S)^2\right) \\ &= \theta_4(P^*_{\theta_3}) - \frac{\gamma^2}{2\delta}(\theta_3 - \bar{\theta}_3^S)^2,\end{aligned} \tag{A.2}$$

where $\theta_4(P^*_{\theta_3})$ denotes end-of-school skills if the child with skills θ_3 is taught at her individually optimal teaching pace $P^*_{\theta_3}$ (we suppress the j-index of P as we consider only one period in this case). Thus, in a given track, end-of-school skills are a strictly decreasing function of the *distance* to the average skill level $\bar{\theta}_3^S$ in that track. This is intuitive given Lemma 1, as it is solely the average skill level to which the instruction pace is optimally targeted.

Next, assume for contradiction that the expected value of end-of-school skills across tracks $\mathbb{E}[\theta_4]$ is maximized under a track allocation mechanism that does not feature a skill threshold. Suppose that $P^V < P^A$ without loss of generality. By Lemma 1, these are the optimal instruction paces for the average skill level in track V and A, respectively. Therefore, $\mathbb{E}(\theta_3|S = V) < \mathbb{E}(\theta_3|S = A)$. Then, because there is no strict threshold, this means that for any initial skill level θ_3, there must be at least two children with initial skill levels smaller or equal to θ_3 that go to different tracks or at least two children with initial skill levels larger or equal than θ_3 that go to different tracks. This implies that there exists a child with $\theta_3' \leq \mathbb{E}(\theta_3|S=V)$ that goes to track $S=A$, and/or a child with $\theta_3' \geq \mathbb{E}(\theta_3|S=A)$ that goes to track $S=V$, and/or two children with skills $\theta_3' < \theta_3''$, with $\theta_3', \theta_3'' \in [\mathbb{E}(\theta_3|S=V), \mathbb{E}(\theta_3|S=A)]$, where the child with the smaller skill level goes to track A and the child with the larger skill level to track V.

However, given the condition in (A.2), this child with θ_3' would always benefit from being in the other track as the distance between her skill level and the average skill level in that track is smaller than in the track she is in. Note that moving just one child to another track does not change the average skills in both tracks. Thus, the policymaker can improve aggregate end-of-school skills by moving this child. The same line of argument holds in the implied game that parents play when they endogenously sort their children into two tracks. If no skill threshold level exists, there is always a child that would unilaterally gain if sent to a different track.

Thus, we have established that the existence of a skill threshold is necessary for optimal end-of-school skills both if a policymaker makes the track allocation directly and when parents play a sorting game. Next, we characterize the thresholds for both cases. Let $\tilde{\theta}_3$ be the skill threshold and let S again indicate to which track a child is allocated, now with $S=V$ for all $\theta_3 \leq \tilde{\theta}_3$ and $S=A$ for all $\theta_3 > \tilde{\theta}_3$.

A policymaker solves

$$\begin{aligned} & \max_{\tilde{\theta}_3} \quad \mathbb{E}(\theta_4) \\ \iff & \max_{\tilde{\theta}_3} \quad \mathbb{E}(\mathbb{E}(\theta_4|S)) \\ & \text{subject to} \\ & \qquad P^S \text{ chosen optimally given Lemma 1.} \end{aligned} \tag{A.3}$$

Using (A.1) and the law of iterated expectations, this maximization problem boils down to

$$\begin{aligned} & \max_{\tilde{\theta}_3} \quad \frac{\beta^2}{2\delta} + \frac{\gamma^2}{2\delta}\,\mathbb{E}\left(\mathbb{E}(\theta_3|S)^2\right) \\ \iff & \max_{\tilde{\theta}_3} \quad \frac{\beta^2}{2\delta} + \frac{\gamma^2}{2\delta}\left(F(\tilde{\theta}_3)\,\mathbb{E}(\theta_3|\theta_3 \leq \tilde{\theta}_3)^2 + (1-F(\tilde{\theta}_3))\,\mathbb{E}(\theta_3|\theta_3 > \tilde{\theta}_3)^2\right), \end{aligned} \tag{A.4}$$

where $F(.)$ denotes the cumulative distribution function of the normal distribution. Note that the right term is just the expected value (across tracks) of the conditional expected values of initial skills squared, conditional on the school track. This corresponds to the variance of the conditional expected values, which depend on the skill threshold $\tilde{\theta}_3$. Using the law of total variance, the maximization problem can thus be rewritten as (dropping the constant term)

$$\begin{aligned} &\max_{\tilde{\theta}_3} \quad \mathbb{E}(\theta_4) \\ \iff &\max_{\tilde{\theta}_3} \quad \frac{\gamma^2}{2\delta}\left(\sigma^2_{\theta_3} - \mathbb{E}(Var[\theta_3|S])\right). \end{aligned} \tag{A.5}$$

Thus, the policymaker chooses optimally a threshold such that the expected variance of skills in each track is minimized. The unique solution is then to set $\tilde{\theta}_3^* = \mathbb{E}\,\theta_3 = 0$, that is, to split the distribution exactly in half. This makes the peer groups in each track as homogeneous as possible, which maximizes average and aggregate learning.

Next, we characterize the threshold that arises endogenously from the sorting game played by the parents. The equilibrium condition maintains that at this threshold, a parent is just indifferent between tracks as her child's skills would be equivalent in both tracks. A parent of a child with skill $\hat{\theta}_3$ is indifferent between tracks V and A iff

$$\begin{aligned} &\left(\alpha + \hat{\theta}_3\frac{\gamma^2}{\delta}\right)\mathbb{E}(\theta_3|\theta_3 \le \hat{\theta}_3) - \frac{\gamma^2}{2\delta}\,\mathbb{E}(\theta_3|\theta_3 \le \hat{\theta}_3)^2 \\ =&\left(\alpha + \hat{\theta}_3\frac{\gamma^2}{\delta}\right)\mathbb{E}(\theta_3|\theta_3 > \hat{\theta}_3) - \frac{\gamma^2}{2\delta}\,\mathbb{E}(\theta_3|\theta_3 > \hat{\theta}_3)^2 \\ \iff &\left(-\alpha - \hat{\theta}_3\frac{\gamma^2}{\delta}\right)\sigma_{\theta_3}\frac{f(\hat{\theta}_3/\sigma)}{F(\hat{\theta}_3/\sigma)} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_3}\frac{f(\hat{\theta}_3/\sigma)^2}{F(\hat{\theta}_3/\sigma)^2} \\ =&\left(\alpha + \hat{\theta}_3\frac{\gamma^2}{\delta}\right)\sigma_{\theta_3}\frac{f(\hat{\theta}_3/\sigma)}{1 - F(\hat{\theta}_3/\sigma)} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_3}\frac{f(\hat{\theta}_3/\sigma)^2}{(1 - F(\hat{\theta}_3/\sigma))^2} \end{aligned} \tag{A.6}$$

where $F(\cdot)$ denotes the CDF of a standard normally distributed random variable, and $f(\cdot)$ is its density function. We solve for the root $\hat{\theta}_3$ that solves (A.6) numerically. In all cases with reasonable parameter values, (A.6) is a monotone function, such that the root is unique if it exists. In the special case without direct peer externality, i.e., $\alpha = 0$, the solution is $\hat{\theta}_3 = 0$, as can be directly seen from (A.6). When $\alpha > 0$, the root is smaller than 0, i.e. $\hat{\theta}_3 < 0$.

Proposition 2

The proof of this Proposition follows directly from (A.5). In a comprehensive system, the variance of initial skills across tracks is just equal to the overall variance since there is only one track. In a tracking system, the expected value of the conditional variances of skills across tracks is smaller than the overall variance, by the law of total variance and provided that the instruction paces are different across tracks. This holds for every skill threshold, not just for the optimal one. Thus average learning is higher.

Next, we show that a full tracking system leads to a "fatter" right tail of the end-of-school skill distribution compared to a comprehensive system. To see this, consider the child who, in expectation, has the highest end-of-school skill in a comprehensive system. Since θ_4 is monotonically increasing in θ_3 in a given track (see (A.1)), this is the child with the highest initial skill, say $\theta_{3,max}$. Moreover, from the properties of a truncated normal distribution, we know that, for any skill threshold $\tilde{\theta}_3$, average skills in the A track, $\bar{\theta}_{3,A}$ are larger than the unconditional average, $\bar{\theta}_{3,C} = 0$. Thus, the squared distance between $\theta_{3,max}$ and $\bar{\theta}_{3,A}$ in a tracking system is smaller. Taken together, (A.2) implies that the child with initial skill $\theta_{3,max}$ ends up with larger end-of-school skills compared to a comprehensive system, which skews the distribution positively.

Finally we derive the range of winners and loser from a tracking system relative to a comprehensive system. Given that θ_4 are monotonically increasing in θ_3 in every track, the range is characterized by the intersection of the linear function $\theta_{4,C}(\theta_3, \bar{\theta}_{3,C})$ with $\theta_{4,V}(\theta_3, \bar{\theta}_{3,V})$ and $\theta_{4,A}(\theta_3, \bar{\theta}_{3,A})$, which are just (A.1) if everyone was taught at the comprehensive, academic, or vocational pace. For any skill threshold, the lower intersection $\theta_{3,L}$ hence solves

$$\begin{aligned} & \theta_{3,L} + \alpha\bar{\theta}_{3,C} + \frac{\beta^2}{2\delta} + \frac{\beta\gamma}{\delta}\theta_{3,L} + \frac{\gamma^2}{\delta}\bar{\theta}_{3,C}\theta_{3,L} - \frac{\gamma^2}{2\delta}\bar{\theta}_{3,C}^2 + \eta_4 \\ = \;& \theta_{3,L} + \alpha\bar{\theta}_{3,V} + \frac{\beta^2}{2\delta} + \frac{\beta\gamma}{\delta}\theta_{3,V} + \frac{\gamma^2}{\delta}\bar{\theta}_{3,V}\theta_{3,L} - \frac{\gamma^2}{2\delta}\bar{\theta}_{3,V}^2 + \eta_4 \\ \iff \quad \theta_{3,L} = \;& \frac{1}{2}\bar{\theta}_{3,V} - \frac{\alpha\delta}{\gamma^2}. \end{aligned} \tag{A.7}$$

Similarly, the upper intersection is given at

$$\theta_{3,U} = \frac{1}{2}\bar{\theta}_{3,A} - \frac{\alpha\delta}{\gamma^2}. \tag{A.8}$$

For any skill threshold $\tilde{\theta}_3$, the interval $[\theta_{3,L}, \bar{\theta}_{3,U}]$ is non-empty. Hence, there are always children with initial skill levels inside this interval who lose in terms of end-of-school skills in a full tracking system relative to a comprehensive system. Every child outside of this interval gains relative to the comprehensive system.

With $\alpha = 0$, the tracking skill threshold is at $\tilde{\theta}_3 = 0$ even if parents endogenously sort their children. Hence, children with initial skills inside a symmetric interval around 0, $[\frac{1}{2}\bar{\theta}_{3,V}, \frac{1}{2}\bar{\theta}_{3,A}]$, lose relative to a comprehensive track, since $\bar{\theta}_{3,V} = -\bar{\theta}_{3,A}$ if $\tilde{\theta}_1 = 0$. The average loss of a child in this interval is equal to $\frac{\gamma^2}{2\delta}\bar{\theta}^2_{3,V} = \frac{\gamma^2}{2\delta}\bar{\theta}^2_{3,A}$.

If $\alpha > 0$, and the policymaker enforces the tracking skill threshold $\tilde{\theta}_3 = 0$, the losses from tracking are concentrated among children in the V track. To see this, note that every child with initial skill in the interval $[\theta_{3,L}, 0]$ is allocated into the V track but loses relative to a comprehensive system. Similarly, every child with an initial skill inside $[0, \theta_{3,U}]$ is allocated to track A but loses relative to a comprehensive system. With $\alpha > 0$, $|\theta_{3,U}| < |\theta_{3,L}|$ and hence, the range of children in the A track that lose is smaller. The interval $[0, \theta_{3,U}]$ may even be empty in which case only children in the V track lose from tracking.

Proposition 3

For the proof of this proposition, we denote by θ_3 the child skills at the beginning of secondary school, by θ_4 the skills at the intermediary stage of secondary school and by θ_5 the skills at the end of secondary school. All other assumptions are maintained. First, we characterize the variance of θ_4. We start by collecting expressions for conditional and unconditional first and second moments.

The unconditional expected value of θ_4 in track V, if everyone went to V is

$$\begin{aligned}\mathbb{E}(\theta_{4,V}) &= \frac{\beta^2}{2\delta} + \alpha\bar{\theta}_{3,V} - \frac{\gamma^2}{2\delta}\bar{\theta}^2_{3,V} \\ &= \frac{\beta^2}{2\delta} - \alpha\sigma_{\theta_1}\frac{f(\tilde{\theta}_3/\sigma_{\theta_3})}{F(\tilde{\theta}_3/\sigma_{\theta_3})} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_3}\frac{f(\tilde{\theta}_3/\sigma_{\theta_3})^2}{F(\tilde{\theta}_3/\sigma_{\theta_3})^2}.\end{aligned} \tag{A.9}$$

The unconditional expected value of θ_4 in track A, if everyone went to A is

$$\begin{aligned}\mathbb{E}(\theta_{4,A}) &= \frac{\beta^2}{2\delta} + \alpha\bar{\theta}_{3,A} - \frac{\gamma^2}{2\delta}\bar{\theta}^2_{3,A} \\ &= \frac{\beta^2}{2\delta} + \alpha\sigma_{\theta_3}\frac{f(\tilde{\theta}_3/\sigma_{\theta_3})}{1 - F(\tilde{\theta}_3/\sigma_{\theta_3})} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_3}\frac{f(\tilde{\theta}_3/\sigma_{\theta_3})^2}{(1 - F(\tilde{\theta}_3/\sigma_{\theta_3}))^2}.\end{aligned} \tag{A.10}$$

The variance of θ_4 in a comprehensive system is

$$\begin{aligned} Var(\theta_{4,C}) &= \mathbb{E}(\big(\theta_4 - \mathbb{E}(\theta_4))^2\big) \\ &= (1+\beta)^2\sigma^2_{\theta_3} + \sigma^2_{\eta_4} \\ &\sigma^2_{\theta_4,C} + \sigma^2_{\eta_4}, \end{aligned} \tag{A.11}$$

where we define $\sigma^2_{\theta_4,C}$ to be the variance of θ_4 net of the additive skill shock variance.

Second, we can derive the expected value of end-of-school skills in the 2-period model in a late tracking system as

$$\begin{aligned} \mathbb{E}(\theta_{5,LT}) &= \mathbb{E}(\mathbb{E}(\theta_{5,LT}|S_{LT})) \\ &= \mathbb{E}(\theta_{4,LT}) + \frac{\beta^2}{2\gamma} + (\alpha+\beta)\,\mathbb{E}(\mathbb{E}(\theta_{4,LT}|S_{LT})) + \frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_{4,LT}|S_{LT})^2) \\ &= (2+\alpha+\beta)\frac{\beta^2}{2\gamma} + \frac{\gamma}{2}[\sigma^{\theta_4,LT} - \mathbb{E}(Var(\theta_{4,LT}|S_{LT}))], \end{aligned} \tag{A.12}$$

where $\mathbb{E}(\theta_{4,LT})$ and $\sigma^2_{\theta_4,LT}$ are just equal to the mean and variance of the comprehensive system in the one-period model (see equation (A.11)). The variable S_{LT} indicates the track selection in period 2, which follows the cut-off rule $S_{LT} = V$ if $\theta_{4,LT} \leq \tilde{\theta}_{4,LT}$ and $S_{LT} = A$ otherwise. The cut-off that maximizes (A.12) is $\tilde{\theta}^*_{4,LT} = \mathbb{E}(\theta_{4,LT}) = \frac{\beta^2}{2\gamma}$. This follows as (A.12) mirrors that of average end-of-school skills in the full tracking system of the one-period model in that average and aggregate $\theta_{5,LT}$ decrease in the expected variance of skills in period 2 across tracks.

Similarly, we find the expected value of end-of-school skills in the 2-period model in an early tracking system as

$$
\begin{aligned}
\mathbb{E}(\theta_{5,ET}) &= \mathbb{E}(\mathbb{E}(\theta_{5,ET}|S_{ET})) \\
&= \frac{\beta^2}{2\gamma} + (1+\alpha+\beta)\,\mathbb{E}(\mathbb{E}(\theta_{4,ET}|S_{ET})) + \beta\frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_{4,ET}|S_{ET})^2) \\
&= \frac{\beta^2}{2\gamma} + (1+\alpha+\beta)\left(\frac{\beta^2}{2\gamma} + \beta\frac{\gamma}{2}[\sigma^2_{\theta_3} - \mathbb{E}(Var(\theta_{3,ET}|S_{ET}))]\right) + \beta\frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_{4,ET}|S_{ET})^2 \\
&= \frac{\beta^2}{2\gamma} + (1+\alpha+\beta)\left(\frac{\beta^2}{2\gamma} + \beta\frac{\gamma}{2}[\sigma^2_{\theta_3} - \mathbb{E}(Var(\theta_{3,ET}|S_{ET}))]\right) \\
&+ \beta\frac{\gamma}{2}[\sigma^2_{\theta_4,ET} - \mathbb{E}(Var(\theta_{4,ET}|S_{ET}))].
\end{aligned}
\tag{A.13}
$$

Comparing (A.12) and (A.13), the condition that governs if average end-of-school skills in a late tracking system are larger than in an early tracking system reads

$$
\begin{aligned}
&\mathbb{E}(\theta_{5,LT}) - \mathbb{E}(\theta_{5,ET}) \\
&= \beta\frac{\gamma}{2}\left(\mathbb{E}(\mathbb{E}(\theta_{4,LT}|S_{LT})^2) - \mathbb{E}(\mathbb{E}(\theta_{4,ET}|S_{ET})^2)\right) \\
&- (1+\alpha+\beta)\beta\frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_3|S_{ET})^2) > 0.
\end{aligned}
\tag{A.14}
$$

The last term of (A.14) represents the advantage of early tracking in the first stage of the schooling years. It stems from the smaller expected conditional variances of initial skills among children that are tracked relative to the overall variance. The conditional expected value of θ_2 in a late tracking system is given by

$$
\mathbb{E}(\theta_{4,LT}|S_{LT} = V) = \frac{\beta^2}{2\gamma} - \sigma_{\theta_4,LT}\frac{f(\tilde{\theta}_{4,LT}/\sigma_{\theta_4,LT})}{F(\tilde{\theta}_{2,LT}/\sigma_{\theta_4,LT})} \tag{A.15}
$$

and

$$
\mathbb{E}(\theta_{4,LT}|S_{LT} = A) = \frac{\beta^2}{2\gamma} + \sigma_{\theta_4,LT}\frac{f(\tilde{\theta}_{4,LT}/\sigma_{\theta_4,LT})}{1 - F(\tilde{\theta}_{4,LT}/\sigma_{\theta_4,LT})}, \tag{A.16}
$$

where the unconditional variance of θ_4 in a late tracking system is given by $\sigma^2_{\theta_4,LT} = \sigma^2_{\theta_4,C} + \sigma^2_{\eta_4}$, i.e. by the one computed in equation (A.11). Since late tracking occurs *after* the realization of skill shocks in period 4, this variance additively *includes* the variance of these shocks.

Condition (A.14) is generally ambiguous and hard to interpret for arbitrary skill thresholds. We focus again on the optimal tracking case, that is, the case with skill threshold $\tilde{\theta}_3 = \mathbb{E}(\theta_3) = 0$ and $\tilde{\theta}_4 = \mathbb{E}(\theta_{4,LT}) = \frac{\beta^2}{2\gamma}$. In that case, we can write the expressions for the various expected square conditional expected values as follows:

$$\mathbb{E}(\mathbb{E}(\theta_3|S_{ET})^2) = 2\chi\sigma_{\theta_3}^2$$
$$\mathbb{E}(\mathbb{E}(\theta_{4,LT}|S_{LT})^2) = \frac{\beta^4}{4\gamma^2} + 2\chi(\sigma_{\theta_{4,LT}}^2 + \sigma_{\eta_4}^2)$$
$$\mathbb{E}(\mathbb{E}(\theta_{4,ET}|S_{ET})^2) = \frac{\beta^4}{4\gamma^2} + 2\chi\sigma_{\theta_3}^2\left(\alpha^2 + \gamma^2 f(0)^2\sigma_{\theta_3}^2 - \frac{\beta^2}{2}\right)$$
$$+2f(0)\sigma_{\theta_3}^2\left(\beta^2 + 2\alpha(1+\beta) - (2\gamma f(0)\sigma_{\theta_3})^2\right) + 2\chi(\sigma_{\theta_{4,LT}}^2 + 2\chi\gamma^2\sigma_{\theta_3}^2).$$

Condition (A.14) then becomes

$$\begin{aligned}
&\mathbb{E}(\theta_{5,LT}) - \mathbb{E}(\theta_{5,ET}) \\
&= \beta\frac{\gamma}{2}\left(2\chi\sigma_{\eta_4}^2 - 2\chi\sigma_{\theta_3}^2\left(\alpha^2 + \gamma^2 f(0)^2\sigma_{\theta_3}^2 - \frac{\beta^2}{2}\right.\right. \\
&\left.\left. + \beta^2 + 2\alpha(1+\beta) - 4\gamma^2 f(0)^2\sigma_{\theta_3}^2 + 2\chi\gamma^2\sigma_{\theta_3}^2 + 1 + \alpha + \beta\right)\right) \\
&= \frac{\gamma}{\pi}\left(\sigma_{\eta_4}^2 - \sigma_{\theta_3}^2\left(1 + \alpha + \alpha^2 + \beta + \frac{\beta^2}{2} + 2\alpha(1+\beta) + \frac{\gamma^2}{2\pi}\sigma_{\theta_3}^2\right)\right) > 0.
\end{aligned} \tag{A.17}$$

From this, Proposition 3 follows.

B Equilibrium Definition

We introduce some notation to define the equilibrium more easily. Let $x_j \in X_j$ be the age-specific state vector of an individual of age j, as defined by the recursive representation of the individual's problems in Section 2. Let its stationary distribution be $\Theta(X)$. Then, a stationary recursive competitive equilibrium for this economy is a collection of: (i) decision rules for college graduation $\{d^E(x_5)\}$, for school track $\{d^S(x_{11})\}$, consumption, labor supply, and assets holdings $\{c_j(x_j), n_j(x_j), a_j(x_j)\}$, and parental transfers $\{a'_5(x_j)\}$; value functions $\{V_j(x_j)\}$; (iii) aggregate capital and labor inputs $\{K, H_0, H_1\}$; (iv) prices $\{r, w^0, w^1\}$; and (v) average skill levels among children in school track S, $\{\bar{\theta'}_{j,S}\}$ such that:

1. Given prices and average skill levels among children in each school track, decision rules solve the respective household problems and $\{V_j(x_j)\}$ are the associated value functions.

2. Given prices, aggregate capital and labor inputs solve the representative firm's problem, i.e. the prices are equal to the marginal products of each input.

3. Given average skill levels among children in each school track, allocation of children in school track solves the parent's problem, i.e. actual average skill levels are consistent with parents' prior.

4. Labor markets for each education level clear.
   For high-school level:

$$H_0 = \sum_{j=5}^{16} \int_{X_j} n_j(x_j)\, h_j\,(x_j)\, d\Theta(X \mid E = 0) + \sum_{j=5}^{5} \int_{X_j} n_j(x_j)\, h_j\,(x_j)\, d\Theta(X \mid E = 1)$$

   where the first summation is the supply of high-school graduates while the second is the labor supply of college students who work during their college education.
   For college level:

$$H_1 = \sum_{j=6}^{16} \int_{X_j} n_j(x_j)\, h_j\,(x_j)\, d\Theta(X \mid E = 1).$$

5. Asset market clears

$$K = \sum_{j=5}^{20} \int_{X_j} a_j(x_j) d\Theta(X),$$

   which implies that the goods market clears;

6. The distribution of X is stationary: $\Theta(X) = \int \Gamma(X) d\Theta(X)$.

C German Education System

In this section, we provide an overview of the most important features of the German Education and School System. A more extensive description can be found, for example, in Henninges et al. (2019) or OECD (2020). Figure C.1 illustrates a simplified structure of the system, starting in Grade 4 and ending with tertiary education.

Generally, schooling is compulsory in Germany for every child starting at age six and lasting until age 18. However, the obligation to go to school typically lasts until grade 9 or 10, after which it shifts to a vocational training obligation if no upper secondary school is attended. At age six, all children visit a comprehensive primary school that lasts the first four grades.[63] After that, children are allocated into traditionally three different secondary school tracks: A lower vocational track, a medium vocational track, and an academic track. However, triggered by the so-called PISA shock in the early 2000s, federal states in Germany have started reforming their secondary school system. In particular, the two vocational tracks have often been combined into one, resulting in a two-track system in the majority of federal states (Bellenberg and Forell, 2012). For that reason, and because even if still two vocational tracks exist, they are much more similar in comparison to the academic track schools, we opt to restrict our analysis in this paper to two school tracks.

Generally, the school tracks differ in the curricula taught, the length of study, and the end-of-school qualifications that come with graduation. In particular, only the academic track schools end with a university entrance qualification that directly allows children to go to college. This requires the completion of the second stage of secondary school, typically grades 10/11 to 12/13. Graduating from a vocational track occurs after Grades 9 and 10 and allows children to take up vocational training in blue-collar jobs or proceed to a professional school that prepares for entry into white-collar, business, or skilled trade occupations. At this stage, there is considerable scope for mobility between tracks. Firstly, professional degrees often allow access to university studies in selected fields. Secondly, children can directly switch to an academic track school if their school marks and achievements admit that. Finally, after having worked for a number of years in vocational jobs, access to some college degrees can be possible. At the same time, it is, of course, possible to switch from an academic track school to a vocational training or job after the mandatory education has been completed.

The public expenditure per student does not differ significantly across school tracks. Table C.1 lists average per-student expenditures across the various school types in the years 2010 to 2020. Across these years, public expenditures by student were highest in pure lower vocational track schools. Expenditures in academic track schools were roughly equal compared to expenditures in joint vocational track schools. The bulk of these expenditures is attributable to teacher pay (around 80%) and the rest for investments into buildings, equipment etc. This suggests that resource differences across school tracks should not be a main driver behind achievement differences, on average.

A remaining driver behind achievement differences across school tracks could be the teaching quality. In particular, higher-quality teachers could select into academic track schools. However, regardless of the secondary school track, becoming a teacher requires

[63] In two federal states, Berlin and Brandenburg, comprehensive primary school lasts the first 6 grades.

Figure C.1: Simplified Structure of the German Education System

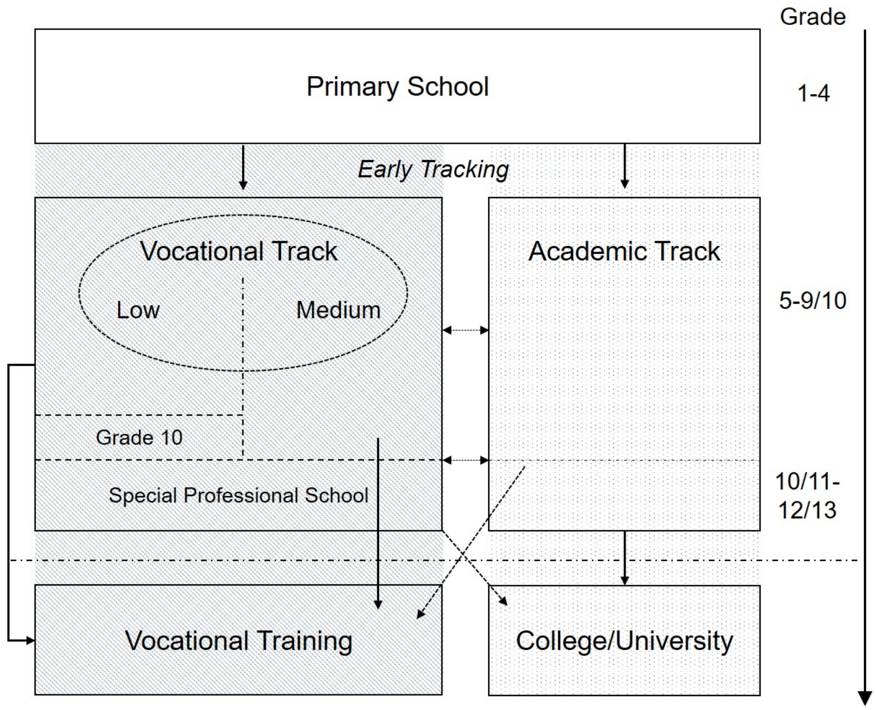


Table C.1: Per-Student Public Expenditures across School Types and Years

Year	Primary	Lower Voc.	Upper Voc.	Joint Voc.	Acad.	Compr.
2010	5,200	7,100	5,300	8,000	6,600	6,600
2011	5,500	7,300	5,600	8,000	7,100	7,100
2012	5,400	7,900	5,700	7,700	7,200	7,200
2013	5,600	8,200	5,900	7,700	7,500	7,500
2014	5,900	8,700	6,200	8,000	7,800	7,800
2015	6,000	8,900	6,400	8,000	7,900	8,000
2016	6,200	9,300	6,700	8,100	8,100	8,200
2017	6,400	9,800	7,000	8,300	8,500	8,600
2018	6,700	10,400	7,400	8,700	8,800	9,100
2019	7,100	11,200	7,900	9,200	9,300	9,500
2020	7,400	12,200	8,200	9,500	9,600	10,000

Source: Statistisches Bundesamt (Bildungsfinanzbericht, Bildungsausgaben - Ausgaben je Schüler, Sonderauswertung)
All amounts in euros.

university studies in the range of 7 to 10 semesters and a similar university degree. On top of that, the differences in wages across school tracks are no longer significant in many federal states. For example, both tenured teachers at vocational track schools and teachers at academic track schools are eligible for the same public pay grade in most northern and eastern federal states already.

D Measuring Child Skills in the NEPS

In this section, we provide an overview of our measures of child skills. One of the main goals of the NEPS project is to document the development of competencies of individuals over their lifespan (Blossfeld et al., 2019). To that end, the NEPS carefully designs and implements regular tests of the respondents' competencies along several domains, including reading comprehension, mathematical competence, and scientific literacy, which we use for the estimation of the child skill technology, but also domains such as information and communication technologies (ICT) literacy. In line with the guidelines set by the Program for International Student Assessment (PISA), the tests are generally designed to assess the extent to which children have learned the content of school curricula but also to judge a child's ability to use domain-specific knowledge to constructively engage with real-life problems (Neumann et al., 2013). The math test, for example, includes items related to "overarching" mathematical content areas that are consistent across all ages, such as quantity, change & relationships, space & shape, as well as several cognitive components, such as mathematical communication, argumentation, or modeling. The age-specific test items include for the majority simple multiple-choice questions with four response options. In addition, the sometimes include more complex multiple-choice questions, as well as short-constructed responses.[64] Each domain is tested using between 20 and 25 items, which usually takes around 30 minutes (Pohl and Carstensen, 2013).

In order to use these questions for the analysis of latent competencies, they need to be scaled. For reading comprehension, mathematical competence, and scientific literacy, the NEPS (similar to the PISA) uses a scaling procedure that is based on item response theory (IRT). IRT is a popular instrument in psychometrics to extract latent ability or other factors from test data. To quote the NEPS: "IRT was chosen as scaling framework for the newly developed tests because it allows for an estimation of item parameters independent of the

[64] A simple multiple choice question consists of one correct out of four answer categories, and complex multiple choice questions consist of a number of subtasks with one correct answer out of two options. Short-constructed responses typically ask for a number (Pohl and Carstensen, 2012).

sample of persons and for an estimation of ability independent of the sample of items. With IRT it is possible to scale the ability of persons in different waves on the same scale, even when different tests were used at each measurement occasion" (Pohl and Carstensen, 2013).

The scaling model used by the NEPS for dichotomous items is the Rasch model (Rasch, 1960).[65] This model assumes that the right answers given to a set of questions by a number of respondents contain all information needed to measure a person's latent ability as well as the question's difficulty. It does so by positing that the probability that person v gives the right answer to question i is given by:

$$p(X_{vi} = 1) = 1 - p(X_{vi} = 0) = \frac{\exp(\theta_v - \sigma_i)}{1 + \exp(\theta_v - \sigma_i)}, \tag{D.1}$$

where θ_v denotes the latent ability of person v and σ_i is a measure of the question's difficulty. Thus, this model maps the total sum score of an individual into an ability parameter estimate. The scale is arbitrary. However, the ability estimate is cardinal.[66] This model is estimated via (weighted) conditional maximum likelihood under a normality assumption on latent ability.

There are several challenges that arise when scaling the test items: These include dealing with different response formats, the treatment of missing responses, adaptive testing, and linking tests across cohorts. An overview about the approaches undertaken by the NEPS to overcome these challenges is given in Pohl and Carstensen (2013). Table D.1 exemplary describes our available NEPS samples of mathematics assessments by starting cohort and grade level.

E Details on Child Skill Technology Estimation

Following the literature on child skill formation, we employ a linear measurement system for the logarithm of latent skills in each period that is given by

$$M_{i,k,j} = \mu_{k,j} + \lambda_{k,j}\theta_{i,j} + \epsilon_{i,k,j}, \tag{E.1}$$

where $M_{i,k,j}$ denotes the kth measure for latent log skills of child i in period j. In each period, we have at least 3 different measures in our data, which typically constitute the achievement (item response theory) test scores of each child in the domains of reading,

[65] For polytomous items, the Partial Credit Model is used, which is a generalization of the Rasch model (Masters, 1982).

[66] It is interval-scaled as Ballou (2009) puts it. That means an increase of 5 points from 15 to 20 represents the same gain in achievement as from 25 to 30.

Table D.1: NEPS Mathematic Assessment Samples

			Information on Parents' Education		Information on School Track	
		Obs.	Obs.	% College Parents	Obs.	% Ac. Track
Cohort 1	K1	2,014	1,709	51%		
Cohort 2	G1	6,352	5,784	46%	2,731	63%
	G2	5,888	5,425	47%	2,651	62%
	G4	6,610	6,068	46%	3,229	63%
	G7	2,479	2,410	51%	2,208	58%
Cohort 3	G5	5,193	3,856	38%	4,369	52%
	G7	6,191	4,214	38%	5,525	49%
	G9	4,888	3,387	38%	4,356	47%
	G12*	3,785	2,830	41%	3,331	58%
Cohort 4	G9	14,523	8,474	35%	14,215	40%
	G12*	5,733	3,767	24%	5,530	23%

Notes: This table describes NEPS mathematics assessments by cohort. Note that in Grade 12, the assessments are different by school track, which makes the comparison of test scores by parental education or school track impossible. Source: NEPS.

maths and scientific literacy. The parameters $\mu_{k,j}$, and $\lambda_{k,j}$ denote the location and factor loading of latent log skills, respectively. By $\epsilon_{i,k,j}$, we denote the measurement error. The parameters and measures are defined conditional on child's age and gender, which we keep implicit.

Following Cunha et al. (2010), we normalize $\mathbb{E}(\theta_j) = 0$ and $\lambda_{1,j} = 1$ for all j. That is, the first-factor loading is normalized to 1 in all periods.[67] We further normalize the measurement errors, such that $E(\epsilon_{k,j}) = 0$ for all j. Given that, the location parameters $\mu_{k,j}$ are identified from the means of the measures. In order to identify the factor loadings, we further assume that the measurement errors are independent of each other across measures and independent from latent skills. Under these assumptions and given that we have at least three measures of latent skills available in each period, we can identify the loadings on each measure in each

[67] We are aware of the potential bias that can arise from this assumption (see Agostinelli and Wiswall (2016)). However, contrary to their case, we measure three different stages of child development, where each stage comes with a new cohort of children (see below). Thus we cannot follow children over multiple periods. Moreover, even if we could, the data we use does not contain age-invariant measures according to their definition.

period by ratios of covariances of the measures (as in Agostinelli et al., 2019):

$$\lambda_{k,j} = \frac{Cov(M_{k,j}, M_{k',j})}{Cov(M_{1,j}, M_{k',j})} \tag{E.2}$$

for all $k, k' > 1$ and $k \neq k'$. Rescaling the measures by their identified location and scale parameters then gives us error-contaminated measures of latent skills for each period as

$$\theta_{i,j} = \frac{M_{i,k,j} - \mu_{k,j}}{\lambda_{k,j}} - \frac{\epsilon_{i,k,j}}{\lambda_{k,j}} = \widetilde{M}_{i,k,j} - \frac{\epsilon_{i,k,j}}{\lambda_{k,j}}. \tag{E.3}$$

Equipped with identified latent variables up to measurement error for all periods, we can plug these into the empirical analogue of the child skill technology (27), which yields

$$\begin{aligned} \widetilde{M}_{i,k,j+1} = \kappa_{0,j} + \kappa_{1,j}\widetilde{M}_{i,k,j} + \kappa_{2,j}\widetilde{M}^2_{i,k,j} + \kappa_{3,j}\overline{\widetilde{M}}_{-i,j,S} \\ + \kappa_{4,j}(\widetilde{M}_{i,k,j} - \overline{\widetilde{M}}_{j,S})^2 + \kappa_{5,j}E_i + \zeta_{i,k,j+1}, \end{aligned} \tag{E.4}$$

where $\overline{\widetilde{M}}_{-i,j,S}$ refers to the average value of the kth transformed measure across all children other than i in a classroom in track S and $\overline{\widetilde{M}}_{j,S}$ to that of the average value of the measures across all children in a school that belongs to track S.

Importantly, the residual $\zeta_{i,k,j+1}$ now contains not only structural skill shocks, $\eta_{i,j+1}$, but also the measurement errors, $\epsilon_{i,k,j}$ as well as interactions of the measurement error with the rescaled measures and even the variance of the measurement errors. For that reason, even if a standard assumption of mean independence of the structural shocks η conditional on all independent variables holds, an OLS estimator of (E.4) will be biased. To account for that, we follow the literature and use Bartlett factors scores to aggregate the different measures into an unbiased score (Agostinelli et al., 2023). As indicated before, we use maths, reading, and science test scores, which we have available across different cohorts and grades (years) in school: We use Cohort 2 for grades 1 and 4, corresponding to the primary school stage in our model (i.e. period $j = 2$); Cohort 3 for grades 5 to 9, which correspond to the first stage of secondary school in the model (i.e. period $j = 3$); and Cohort 4 for grades 9 to 12, corresponding to the second stage of secondary school in the model (i.e. period $j = 4$). Note that in grade 7, children only take two tests, which is why we cannot construct the latent skills. In addition, in grade 12, the maths test differs by track, and only children in the academic track take the science test. Consequently, in grade 12, we can only create latent skills for children in the academic school track. [68]

Table E.1 summarizes the estimated coefficients of the child skill technology (27) using the identified latent variables as describes above in columns (1) and (2), or using math test scores directly in columns (3) and (4). The estimates differ slightly depending on whether we use longitudinal weights or not, but overall are quite consistent. Table E.2 performs the estimation where the squared distance to track average term in (27) is distributed, such that we include directly the interaction between own skill and track average. The estimated coefficient is positive, statistically significant in most specifications and not statistically different from $-2\hat{\omega}_4$, lending support to our modeling assumptions.

F Empirical Evidence on School Track Selection

In this section, we present reduced-form evidence on the effect of parental background on the school track choice for their children.

Table F.1 shows that parents frequently deviate from teacher recommendations toward their own education. Research on school tracking has found that parents with higher socioeconomic status are more likely to send their child to an academic track school than parents with a lower socioeconomic status, even conditional on school performance or achievement test scores before the track decision. Consistently, we find that 54% of children from college-graduated parents receive a teacher recommendation for the academic track versus 39% of children from non-college-graduated parents.[69] In addition, Table F.1 shows that while around 23% of parents who themselves have a college education overrule a vocational recommendation, only 4% of them overrule an academic recommendation. At the same time, while 16% of non-college graduated parents overrule an academic recommendation, only 12% of them overrule a vocational recommendation. As argued before, one reason for these deviations may be that parents may have more information about their child's skills than teachers. However, the deviations are not symmetric across tracks, and parents are more

[68] In Germany, the vocational track schools typically end after grade 9 or grade 10 and so-called upper secondary schooling only happens in academic track schools. However, the NEPS data keeps track of the students even if they are no longer enrolled in a school and tests them at the same age. A remaining issue is, of course, that even though we know the classroom compositions in grade 9, we do not know how long learning in that classroom continues in a vocational track school. For that reason, we make the assumption that children who went to a vocational track school that finished before they are 18 years old continue to learn in an environment that is the same as if the vocational school had continued. In reality, students who graduate from vocational schools often continue with an apprenticeship, where we think it reasonable to assume that the peer composition is similar to the one in school.

[69] We define children from college parents if they have at least one of the parents with a college education.

Table E.1: Robustness Checks: Child Skill Technology Parameters Estimates

Grade 9 on Grade 5		Skills		Math scores	
Dependent Variable: $\theta_{i,j+1}$		(1)	(2)	(3)	(4)
$\hat{\omega}_{1,3}$	$\theta_{i,j}$	0.664***	0.647***	0.519***	0.517***
		(0.022)	(0.025)	(0.025)	(0.030)
$\hat{\omega}_2$	$\bar{\theta}_{-i,j,S}$	0.003	0.028	0.022	0.025
		(0.020)	(0.021)	(0.024)	(0.031)
$\hat{\omega}_3$	$\theta_{i,j}^2$	0.008*	0.006	0.010**	0.015**
		(0.004)	(0.005)	(0.005)	(0.006)
$\hat{\omega}_4$	$(\theta_{i,j} - \bar{\theta}_{j,S})^2$	-0.011*	-0.013**	-0.012*	-0.020**
		(0.006)	(0.006)	(0.007)	(0.008)
$\hat{\omega}_{5,3}$	$E = 1$	0.034***	0.033***	0.033***	0.045***
		(0.010)	(0.012)	(0.012)	(0.014)
Obs.		1,847	1,676	2,084	1,708
Weights		No	Yes	No	Yes

Notes: This table presents the coefficients of regressions of skills in grade 9 on skills in grade 5, skills squared, the average skill level of peers, distance to the average skill in the track squared, and parent's education dummy. In Columns (2) and (4), all observations are weighted using longitudinal weights, while in Columns (1) and (3), they are not. Standard errors are clustered at the school level. Columns (1) and (2) present the results for latent skills corrected for measurement errors, while columns (3) and (4) present the results for uncorrected latent skills of maths grades. Models control for year of birth, gender, and school-fixed effects. Source: NEPS.

Table E.2: Robustness Checks: Alternative Child Skill Technology Parameters Estimates

Grade 9 on Grade 5		Skills		Math scores	
Dependent Variable: $\theta_{i,j+1}$		(1)	(2)	(3)	(4)
$\hat{\omega}_{1,3}$	$\theta_{i,j}$	0.657***	0.626***	0.515***	0.505***
		(0.021)	(0.024)	(0.023)	(0.028)
$\hat{\omega}_2$	$\bar{\theta}_{-i,j,S}$	0.001	0.024	0.020	0.018
		(0.020)	(0.021)	(0.024)	(0.030)
$-2 * \hat{\omega}_4$	$\theta_{i,j} * \bar{\theta}_{j,S}$	0.018**	0.014	0.022**	0.029**
		(0.009)	(0.010)	(0.010)	(0.012)
$\hat{\omega}_{5,3}$	$E = 1$	0.034***	0.034***	0.033***	0.048***
		(0.010)	(0.012)	(0.012)	(0.014)
Obs.		1,847	1,676	2,084	1,708
Control for $\bar{\theta}^2_{j,S}$		Yes	Yes	Yes	Yes
Weights		No	Yes	No	Yes

Notes: This table presents the coefficients of regressions of skills in grade 9 on skills in grade 5, the average skill level of peers, the interaction between child skills and the average skill in the track, the average skill in the track squared, and parent's education dummy. In Columns (2) and (4), all observations are weighted using longitudinal weights, while in Columns (1) and (3), they are not. Standard errors are clustered at the school level. Columns (1) and (2) present the results for latent skills corrected for measurement errors, while columns (3) and (4) present the results for uncorrected latent skills of maths grades. Models control for year of birth, gender, and school-fixed effects. Source: NEPS.

Table F.1: School Track Choice

Recommendation	Shares	% deviate	% in the top 25% in G9: if followed	% in the top 25% in G9: if deviated
College Parents				
Academic	56%	4%	34%	68%
Vocational	44%	23%	34%	6%
Non-college Parents				
Academic	38%	16%	21%	44%
Vocational	62%	12%	22%	14%

Notes: This table provides information on school track choice by parental education and teacher recommendation. Source: NEPS, Cohort 3.

likely to deviate from teachers' recommendations for their own education.

Parents may have several reasons for frequently overruling teachers' recommendations when they differ from their own education. For instance, they may be better equipped to support their child in a track with which they are more familiar. However, the last columns of Table F.1 show that children of college-educated parents who deviate from the recommended vocational track do relatively poorly compared to those who received the academic recommendation. In fact, only 6% of children of college-educated parents who deviated from the vocational track recommendation belong to the top quartile of skills four years later in Grade 9. In contrast, the same number reaches 34% among those who received an academic track recommendation. This suggests that the support provided by college-educated parents does not fully compensate for relatively low skill levels. Conversely, children from non-college-educated parents who deviate toward the vocational track do remarkably well in Grade 9, with almost half of them belonging to the top quartile of skills. As a comparison, 22% of those with a vocational recommendation reached the top quartile in Grade 9. Those numbers indicate that these students might have succeeded in the academic track as well. Thus, we argue that the relatively high number of deviations towards parents' education is partly driven by a parental bias towards their own education, which is not only motivated by parents' ability to support the child or their intrinsic knowledge of their skills.

G Details on the Data Moments used in the MSM Estimation

In this section, we present details of the data moments that we use as calibration targets in the method of simluated moments estimation.

Table G.1 presents the distribution of students across school tracks and education levels. We use two main sources to compute those shares. First, whenever available, we use official statistics that are reported in the education report: 44% of students are in the academic track, and children from college parents are 2.27 more likely to graduate from college than students from non-college parents.[70] Second, we complement this data using Cohort 4 of the NEPS dataset: 35% of the parents graduated from college, children from college parents are 1.95 more likely to attend the academic school track than children from non-college parents and students in the academic track are 5.23 more likely to attend college than students in the vocational track.[71] In addition, 23% of college parents deviate from the vocational recommendation, and 16% of non-college parents deviate from the academic recommendation as argued above. Finally, the model is in stationary equilibrium, which implies that 35% of students graduate from college, the same share as the share of college parents. All the remaining shares are computed so that the model distribution is internally consistent.

Table G.2 describes the evolution of child skills over time and across groups using the identified latent variable (Columns (1) and (2)) or maths scores directly (Columns (3) and (4)). As before, we use different cohorts of NEPS for the estimation: Cohort 2 for grades 1 to 4, Cohort 3 for grades 5 to 9, and Cohort 4 for grades 9 to 12. We also report the results for grade 12 using Cohort 3, but we prefer the results from Cohort 4 as the number of observations is greater (see Table D.1). For a given individual, the correlation across skills increases over time, from 0.61 between grades 1 and 4 to 0.74 between grades 9 and 12 (Table G.2 Column (1), using the latent skill and longitudinal weights). The differences in average skills across groups are also increasing over time: from 0.541 SD in grade 1 to 0.677 SD in grade 9 by parents' education, and from 0.847 SD in grade 1 to 1.036 SD in grade 9 by school track (Table G.2 Column (1), using the latent skill and longitudinal weights).[72]

Tables G.3 and G.4 report details on the estimation of academic school track attendance

[70] https://www.bpb.de/system/files/dokument_pdf/bpb_abb8_Hochschulbildung_Bildungsherkunft.pdf

[71] In the NEPS dataset, we only have college attendance and not graduation. We use the ratio of college attendance by groups as a proxy for the ratio of college graduation.

[72] To compute the difference by school track in grade 1, we use the panel structure of NEPS, and allocate students in grade 1 to school track according to their actual school track in grade 7.

Table G.1: Distribution of Students, School Tracks and Parental Education

Statistics	Value	Source	Comment
% of college parents	35%	NEPS Cohort 4	
Track choice			
% in ac. track	44%	Education report p.110	42% in NEPS Cohort 4
Ratio % ac. track if college parents to % if non-college parents	2.06	NEPS Cohort 4	
% in ac. track if $E = 1$	66%	Implied	
% in ac. track if $E = 0$	32%	Implied	
Track recommendation			
Deviation if recom. S = 0 and $E = 1$	23%	NEPS Cohort 4	
Deviation if recom. S = 1 and $E = 0$	16%	NEPS Cohort 4	
% ac. recom.	44%	Implied	
% ac. recom. if $E = 1$	56%	Implied	
% ac. recom. if $E = 0$	38%	Implied	
College graduation			
% who graduate from college	35%	Model assumption	
Ratio % college if academics to % if vocational	6.27	NEPS Cohort 4	
% college if academics	66%	Implied	
% college if vocational	11%	Implied	
Ratio % college if college parents to % if non-college parents	3.20	Stifterverband für die Deutsche Wissenschaft e.V. (2021): Vom Arbeiterkind zum Doktor.*	Ratio computed from Figure 1, where 64% of children from college parents are bachelor graduates versus 20% of children from non-college parents.
% college if college parents	63%	Implied	
% college if non-college parents	20%	Implied	

Notes: This table provides information on the distribution of students by school track, college education, and parental education with corresponding sources.
* `https://www.stifterverband.org/medien/vom_arbeiterkind_zum_doktor`

Table G.2: Evolution of Skills

	Skills		Math grades		
Statistics	(1)	(2)	(3)	(4)	Source
Group Differences					
Differences in average skills by parental education (in standard deviations)					
Grade 1	0.530	0.541	0.459	0.462	NEPS Cohort 2
Grade 5	0.658	0.647	0.605	0.579	NEPS Cohort 3
Grade 9	0.672	0.774	0.598	0.697	NEPS Cohort 3
Grade 9	0.710	0.677	0.659	0.623	NEPS Cohort 4
Differences in average skills by school track (in standard deviation)					
Grade 1	0.840	0.847	0.767	0.769	NEPS Cohort 2
Grade 5	1.104	1.022	1.067	0.986	NEPS Cohort 3
Grade 9	1.058	1.089	1.040	1.113	NEPS Cohort 3
Grade 9	1.110	1.036	1.062	0.998	NEPS Cohort 4
Rank-Rank correlations					
Panel A: All students					
Grades 1 to 4	0.72	0.72	0.59	0.58	NEPS Cohort 2
Grades 5 to 9	0.79	0.79	0.71	0.71	NEPS Cohort 3
Panel B: Academic students					
Grades 1 to 4	0.62	0.61	0.46	0.45	NEPS Cohort 2
Grades 5 to 9	0.68	0.69	0.57	0.59	NEPS Cohort 3
Grades 9 to 12	0.74	0.72	0.65	0.66	NEPS Cohort 3
Grades 9 to 12	0.74	0.74	0.66	0.59	NEPS Cohort 4
Panel C: Vocational Students					
Grades 1 to 4	0.64	0.64	0.53	0.50	NEPS Cohort 2
Grades 5 to 9	0.74	0.75	0.63	0.64	NEPS Cohort 3
Weights	No	Yes	No	Yes	

Notes: This table provides information on average differences in skills in one standard deviation unit by parental background and school track over time as well as skill rank-rank correlations. In columns (2) and (4), all observations are weighted with longitudinal weights, while in columns (1) and (3), they are not. Columns (1) and (2) present the results for latent skills corrected for measurement errors, while columns (3) and (4) present the results for uncorrected latent skills of maths grades. Sources are mentioned in the last column.

on child skills at the beginning of secondary, or end of primary school, as well as the estimation of college attendance on past skills and school track. We use the latter estimates to calibrate the college costs in our model, while the former serve as untargeted tests.

Table G.3: School Track on Past Skills

Dependent Variable: Academic School Track	
Panel A: Cohort 3 - Grade 5	
$\theta_{i,j-1}$	0.877***
	(0.019)
Obs	3,888
Panel B: Cohort 2 - Grade 4	
$\theta_{i,j-1}$	0.745***
	(0.027)
Obs	2,299

Notes: This table presents the coefficients of regressions of the academic school track on past skills in grade 4 (Panel A) or in grade 5 (Panel B). Models control for year of birth and gender fixed effects. Source: NEPS.

Table G.4: College on Past Skills and School Track

Dependent Variable: College Attendance	
Panel: Cohort 4 - Grade 9	
$\theta_{i,j}$	0.395***
	(0.015)
S	0.407***
	(0.011)
Obs	10,074
Variance of residuals	0.137

Notes: This table presents the coefficients of regressions of college attendance on past skills (grade 9) and school track. We control for year of birth and gender fixed effects. Source: NEPS.

H Discussion on Child Skill Shocks

As for the adult human capital, we assume child skills are subject to idiosyncratic shocks. These shocks represent unexpected heterogeneity in child development speeds (such as late-bloomers) and any shock that can arise during childhood and affect the child's learning, such as health issues, a move, parents' divorce, meeting an influential mentor, etc.

An alternative model would assume child skills are not subject to shocks but imperfectly observed by parents. In this section, we elaborate on an alternative model based on our baseline model that introduces this feature and compare it to our baseline model.

Specifically, in this alternative modeling, θ would be the true (log) skills that matter for the child skill evolution and future earnings and evolve according to the stage-specific function f, defined by:

$$\theta_{j+1} = f(\theta_j, P_j^S, \bar{\theta}_j^S, E) \tag{H.1}$$

$$= \kappa\theta_j + \alpha\bar{\theta}_j^S + + \beta P_j^S + \gamma\, \theta_j\, P_j^S - \frac{\delta}{2}\, P_j^{S2} + \zeta E, \tag{H.2}$$

where, similarly to the baseline model, P^S is the instruction pace in track S, the average peer skills is denoted by $\bar{\theta}^S$ and E stands for parental background. However, in this alternative version, parents would not directly observe their child's skills θ_j. Instead, in every period, they would receive an unbiased signal $\hat{\theta}_j$ about their child skills, with:

$$\begin{aligned} \hat{\theta}_j &= \theta_j + \epsilon_{\theta,j} \\ \epsilon_{\theta,j} &\sim \mathcal{N}(0, \sigma^2_{\epsilon_\theta}). \end{aligned} \tag{H.3}$$

Given the parents' initial prior $\tilde{\theta}_{j-1}$, that is unbiased and follows a normal distribution $\mathcal{N}(\theta_{j-1}, \sigma_{j-1})$, parents update their perception of their current child's skills $\tilde{\theta}_j^P = f(\tilde{\theta}_{j-1}, P_j^S, \bar{\theta}_j^S, E)$ using Bayesian updating:[73]

[73]We could assume the first initial prior to be equal to the signal they receive in $j = 1$.

$$
\begin{aligned}
\tilde{\theta}_j &= k\ \hat{\theta}_j + (1-k)\tilde{\theta}_j^P \\
\sigma_j^2 &= \sigma_{j-1}^2 - k\sigma_{j-1}^2 \\
k &= \frac{\sigma_{j-1}^2}{\sigma_{j-1}^2 + \sigma_{\epsilon_\theta}^2},
\end{aligned}
\tag{H.4}
$$

where k is the Kalman gain and is increasing in the precision of the signal $(\frac{1}{\sigma^2_{\epsilon_\theta}})$.

Since the perception of child skills is unbiased, the perception of the peer skills is equal to the truth in the limit. Consequently, $\bar{\theta}_j^S$ is assumed to be perfectly observed by the parents and stable in equilibrium. Similarly, in the limit, the policymaker perfectly observed the average child skills in every school track and set the pace of instruction P_j^S according to Lemma 1. Then, we can define the child skill production function as

$$
\begin{aligned}
\theta_{j+1} &= f(\theta_j, \bar{\theta}_j^S, E) \\
&= \frac{\beta^2}{2\delta} +\ (\kappa + \frac{\beta\gamma}{\delta})\theta_j +\ (\alpha)\bar{\theta}^S -\ \frac{\gamma^2}{2\delta}\bar{\theta}^{S2} +\ \frac{\gamma^2}{\delta}\theta_j\bar{\theta}^S + \zeta\ E \\
&= \omega_0 +\ \omega_1\ \theta_j +\ \omega_2\ \bar{\theta}^S +\ \omega_4\ \bar{\theta}^{S^2} -\ 2\omega_4\ \theta_j\bar{\theta}^S +\ \omega_5\ E.
\end{aligned}
$$

Notice that the child skill evolution is identical to one in the baseline model but for the idiosyncratic shock η that are here absent. As a result, the average skill threshold that determines the school track allocation would be identically determined in both model versions. Indeed, in the baseline model, the expected future child skills are independent of the shocks η that are assumed to be normally distributed and centered to zero. To see this, notice that in both models, the average skill threshold θ^* for a given parental background E and current (perceived) skills θ_3, is determined by the following equation:

$$
\begin{aligned}
E(\theta_5, E'|S = A, E) &= E(\theta_5, E'|S = V, E) \\
E(f(\theta_4, \bar{\theta}_4^A, E), E'|E) &= E(f(\theta_4, \bar{\theta}_4^V, E), E'|E) \\
E(\omega_1\ \theta_4 +\ \omega_2\ \bar{\theta}_4^A +\ \omega_4\ \bar{\theta}_4^{A^2} -\ 2\omega_4\ \theta_4\bar{\theta}_4^A, E'|E) &= E(\omega_1\ \theta_4 +\ \omega_2\ \bar{\theta}_4^V +\ \omega_4\ \bar{\theta}_4^{V^2} -\ 2\omega_4\ \theta_4\bar{\theta}_4^V, E'|E).
\end{aligned}
$$

Assuming $\bar{\theta}_j^A$ and $\bar{\theta}_j^V$ for $j = 3, 4$ are known and fixed, by the linearity of the function, we can replace θ_4 in the expectation by its expected value $E(\theta_4|S, E) = f(\theta_3, \bar{\theta}_3^S, E)$. So θ^* is independent of η in the baseline model and identically determined as in this alternative model.

Conceptually, misallocation sources, however, differ between the two models. In the alternative model, at the time of the school track choice $j = 3$, parents make their decision based on their perception of their child's skills $\tilde{\theta}_3 \sim \mathcal{N}(\theta_3, \sigma_3^2)$. Part of the misallocation will be driven by σ_3, which governs how imprecise the parental perception of the skills is. In the baseline model, parents perfectly observed their current child's skills, but skills are subject to shocks. Part of the misallocation is then governed by $\eta_4 \sim \mathcal{N}(0, \sigma_{\eta_4}^2)$, and more precisely by its variance σ_{η_4}. While allowing for re-tracking would solve the issue of misallocation driven by skill uncertainty in the baseline model, it would not completely solve the issue driven by imperfectly observed skills in the alternative model. Indeed, skills are still imprecisely observed in period 4—even though the precision is greater than in period three due to the learning process.

Finally and crucially, earnings variance would be entirely determined at the earliest age without child skills shocks. As a result, comparing early and late tracking in the two model versions leads to different results. While in the baseline model, late tracking versus early tracking makes the school track choice less dependent on early skill conditions, it is the reverse in the alternative model. Postponing tracking allows parents to make a more informed decision about the school track choice, strengthening the relationship between early (true) skill conditions and the school track. Still, the effect on mobility is ambiguous as late tracking shrinks the difference in skills across socioeconomic groups.

In reality, it is probably a mix of both modeling versions. However, the data does not allow us to differentiate between the two mechanisms. We use latent skills for calibration purposes and don't have information on parents' perceptions of their child's skills. Since we think skills are likely subject to shocks during childhood, as human capital is likely subject to shocks during adulthood, we favor the modelization with child skill shocks. The noise in the preference shifter can be regarded as a reduced form of capturing the imprecision in the parents' perception of their child's skills.

I Welfare Measure

Our analysis centers on evaluating aggregate welfare under scenarios that feature different policies. Welfare is defined by the consumption equivalence under the veil of ignorance in the baseline economy relative to the economy with the counterfactual policy in place. Formally, let $C \in \{0, 1, 2, ...\}$ denote the set of counterfactuals, with C = 0 being the baseline economy (early tracking) in a steady state. We refer to the consumption equivalence as the percentage change in consumption Δ in the baseline economy that makes individuals indifferent between being born in the baseline economy (C = 0) and the one in which the counterfactual policy $C \neq 0$ is in place. Denote $\mathcal{V}_5^C(\theta_5, a_5, \phi, S, E^p, \Delta)$ be the welfare of agents with the initial state of the economy j=5 if their consumption (and that of their descendants) were multiplied by $(1+\Delta)$:

$$\mathcal{V}^C(\theta_5, a_5, \phi, S, E^p, \Delta) = \mathbb{E}^C \sum_{j=5}^{j=20} \beta^{j-5} v_j\left(c_j^{*C}(1+\Delta), n_j^{*C}, E^{*C}, \theta_5, S, E^p\right) + \beta^{13-5}\delta \mathcal{V}_{j^5}^C\left(\theta_5', a_5', \phi', S', E^{*C}, \Delta\right),$$

where E^p is the education of the parent, and for j = 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20

$$v_j(c_j, n_j, E, \theta_5, S, E^p) = \frac{(c_j/q)^{1-\sigma}}{1-\sigma} - b\frac{n_j^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}, \tag{I.1}$$

for j = 5

$$v_j(c_j, n_j, E, \theta_5, S, E^p) = \frac{(c_j/q)^{1-\sigma}}{1-\sigma} - b\frac{n_j^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} - 1\{E=1\}\ \psi(S, \theta_5, \nu(E^p)), \tag{I.2}$$

for j = 11

$$v_j(c_j, n_j, E, \theta_5, S, E^p) = \frac{(c_j/q)^{1-\sigma}}{1-\sigma} - b\frac{n_j^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} - 1\{S=A\}\ \chi(E). \tag{I.3}$$

Note that the policy functions are assumed to be unchanged when Δ is introduced. The average welfare is:

$$\bar{\mathcal{V}}^C(\Delta) = \sum_{S,E^p} \int_{\theta_5, a_5, \phi} \mathcal{V}^C(\theta_5, a_5, \phi, S, E^p, \Delta)\mu_C(\theta_5, a_5, \phi, S, E^p)$$

where μ_C is the distribution of initial states $\{\theta_5, a_5, \phi, S, E^p\}$ in the economy C.

We define Δ^C as the consumption equivalence that makes individuals indifferent between being born in the baseline economy $C = 0$ and one in which policy $C \neq 0$ is in place, such that:

$$\bar{\mathcal{V}}^0(\Delta^C) = \bar{\mathcal{V}}^C(0)$$