# Efficiency and Equity of Education Tracking

## A Quantitative Analysis*

Suzanne Bellue†

Lukas Mahler ‡

June 12, 2023

**Abstract**

We study the long-run aggregate and distributional 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 ingredient in the model is the child skill production technology, where a child's skill development depends on her classroom peers and the instruction pace in her school track. We show analytically that this technology can rationalize reduced-form evidence on the effects of school tracking on the distribution of child skills. We calibrate the model using representative data from Germany, a country with a very early and strict school tracking policy. Our model suggests that eliminating the parental influence on the school track choice that arises purely from own-track preferences improves both social mobility and aggregate economic output. An education reform that postpones the tracking age from ten to fourteen generates similar improvements in intergenerational mobility. However, these come at the cost of modest efficiency losses in aggregate economic output. The size of these losses depends on the design of the instruction levels in each school track and on the presence of general equilibrium effects in the labor market.

*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 the seminar and conference participants at the University of Mannheim 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.

†University of Mannheim (suzanne.bellue@uni-mannheim.de)

‡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.[1] The argument behind tracking is typically one of efficiency: grouping children according to their ability and aspirations creates more homogeneous peer groups and allows for tailored instruction levels and curricula, which improve the educational outcomes of children (Bonesrønning et al., 2022; Duflo et al., 2011). However, because the track decision is, in reality, often influenced by family background, it has been suggested that school tracking policies hamper equality of opportunity in access to education and thus impair mobility in educational and labor market outcomes across generations (Falk et al., 2020; Dustmann, 2004; Meghir and Palme, 2005; Pekkala Kerr et al., 2013). The parental influence on the track decision seems particularly strong if tracking occurs at a relatively young age, when measures of child ability are especially noisy (Hanushek and Wössmann, 2006). For that reason, it is not surprising that school tracking, and in particular its timing, is a frequently recurrent issue in the public and academic debate about education reforms in countries with a strict and early tracking regime, such as Germany.[2]

In this paper, we contribute to that debate by providing a quantitative assessment of the long-run aggregate, distributional and inter-generational effects of school tracking policies. In light of the arguments above, any such assessment needs to take into account the effects of tracking on the educational outcomes of children, as well as how these outcomes translate into labor market outcomes and outcomes across generations. This is hard, if not impossible, to do in a purely reduced-form way, not only because of its demands on data, but also because a change in the allocation of children across tracks and, consequently, a change in the allocation of workers across skill levels, may entail general equilibrium effects.[3] Macroeconomic models of mobility provide a useful environment to consider such effects but have so far largely ignored how the development of child skills during school is affected through peers and teaching levels. We fill this gap by building a macroeconomic model of

[1] An overview about school tracking policies in OECD countries is given in Chapter 2 in OECD (2013). We differentiate school tracking, which refers to the allocation of 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.

[2] There is substantial variation in the timing of tracking across OECD countries. While in some countries, such as Germany and Austria, tracking occurs already at the age of 10, in other countries, like the US and UK, do not track at all during secondary school.

[3] For example, if the share of children who are allocated to an academic track school increases substantially, in the long run, the price of academically skilled labor in the economy should decrease. This, in turn, makes an academic track school less attractive, which affects the share again.

overlapping generations that explicitly zooms in on the schooling years of children.

The model is built around a parsimonious theory of how a child's skills are developed during school years. Going to a school that belongs to a particular school track affects child skills directly through interactions with peers at her school and the pace of instruction that is taught in that school. Every child is assumed to have an ideal instruction pace at which she learns best. However, there can only be one instruction pace per school track, which is set endogenously by the policymaker. We show analytically that, under linear direct peer effects and complementarity assumptions between own skill and instruction pace, this gives rise to efficiency gains from tracking in terms of improving aggregate end-of-school skills. Indeed, absent any unforeseeable shocks to child skills, an optimal tracking policy should perfectly stratify children according to their skills as early as possible. In the presence of skill shocks, however, it can be optimal to postpone tracking, even from an efficiency point of view. Moreover, we show that tracking can increase overall inequality in educational outcomes relative to a comprehensive school system. Finally, the theory implies that not all children gain from tracking and that the losses are often concentrated in the track with the lower average skill level. Thus, our child skill formation technology rationalizes some of the most robust empirical findings regarding school tracking in the literature and embeds the main arguments about school tracking that are frequently made in the public discourse.[4]

We embed this child skill development theory into a full general equilibrium life-cycle Aiyagari framework of overlapping generations, in which parents care about their offspring in the tradition of Becker and Tomes (1986) and child skills during the school years evolve according to our technology and are subject to uninsurable skill shocks. The model is tailored to fit the German Education System, where children are tracked into two school tracks at the age of 10 based on a decision by the parents. As in the data, the track decision may be influenced by parental preferences for children to follow in their own education steps. While only one track directly facilitates access to college education, we allow for second-chance opportunities as children can decide to switch tracks after secondary school. Going to college 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 tertiary education decision, the labor earnings. The distribution of human capital across college and non-college workers affects wages, which in turn affects the school track choice.

[4]For the case of Germany, see for instance Matthewes (2021) who shows that earlier tracking raises inequality in educational outcomes and Piopiunik (2014), who shows that low-achievers may be negatively affected by school tracking.

Finally, households can save into a non-state-contingent asset subject to life-cycle borrowing constraints. When children become independent, parents can also make a non-negative inter-vivos transfer.

The model is solved numerically, and the parameters are calibrated in two steps. First, we estimate the parameters of the child skill formation technology directly from German data on school children using a latent variable framework as in Cunha et al. (2010). In particular, we use information on achievement test scores as measures of 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 validity of the model, we investigate 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. Dustmann et al. (2017) argue that for such marginal children in Germany, the initial track choice is inconsequential for labor earnings later in life. Simulated data from our model confirms that children who go to different school tracks solely based on small differences in skills at the time of the track decision experience very similar lifetime economic outcomes.

Notwithstanding this, our quantitative results show that for lifetime inequality across the population, skill formation during the school tracking years, and hence the school tracking policy plays an important role. In particular, variation in the initial school track alone can account for 12% of the variation in lifetime earnings and 13% of the variation in lifetime wealth. As in the data, parental education is, after child skills, the second most important determinant of initial school track choice. This effect comes from direct parental preferences for children to follow in their own track, but also college tastes and knowledge about the deterministic influence of parental education on child skill development. Biased parental preferences in school track choice give rise to inefficiencies in the allocation of children across tracks. For example, a college-educated parent may push her child into an academic-track school even though her child's skills would optimally suggest a vocational-track school. This harms not only her own child's learning outcomes but also affects average learning in that track as the instruction pace endogenously adjusts to the composition of skills in that track.

Using our calibrated model, we perform counterfactual experiments that eliminate the parental bias in school track choice either directly or indirectly by introducing a strict skill threshold that governs school track allocation. The former policy improves social mobility across generations as the intergenerational income elasticity decreases by around 2%, while increasing aggregate output by 0.1%, as learning during the school years becomes more efficient. The latter policy increases aggregate output even further but does not benefit social mobility.

Finally, we use our model to study the long-run effects of an education reform that universally postpones the school tracking age by four years. Such a reform is often suggested in countries with traditionally early tracking systems, such as Germany, as a means to improve equality of opportunity in access to academic education (Woessmann, 2013). We show that postponing the tracking age indeed improves social mobility as it leads to a decrease in the intergenerational elasticity of income of around 2%. These mobility gains arise because the later school track choice is less dependent on the parental background, but also because college education after secondary school is significantly less driven by the previous school track. However, postponing tracking comes at the cost of a 0.2% drop in GDP. The reason for this is that prolonged learning in a comprehensive school track foregoes efficiency gains from tailored instruction levels in an early tracking system. These learning losses cannot be recuperated by efficiency gains coming from the fact that the late tracking decision is based on more complete information about children's skills. We also highlight the importance of considering general equilibrium effects on the labor market that influence school track decisions. In particular, the aggregate output losses in partial equilibrium would be significantly higher at around 0.8-1% of GDP.

## Related Literature

This paper links several strands of the literature: the quantitative family-macroeconomics literature, the children's skill formation literature, and the school tracking literature.

First, we contribute to the quantitative family macroeconomics literature that studies determinants of the intergenerational transmission of economic status (Abbott et al., 2019; Caucutt and Lochner, 2020; Daruich, 2022; Jang and Yum, 2022; Fuchs-Schündeln et al., 2022; Fujimoto et al., 2023; Lee and Seshadri, 2019; Yum, 2022). Some of these studies incorporate a part of the educational system into their analysis, such as Abbott et al. (2019); Caucutt and Lochner (2020); Fuchs-Schündeln et al. (2022) who model high-school graduation choice. However, all of these studies except Fujimoto et al. (2023) focus on the United States, often concentrating on access to higher education and neglecting the importance of designing the (secondary) school system for macroeconomic outcomes. We explicitly focus on the secondary schooling system. Our paper is perhaps most closely related to to Fujimoto et al. (2023) who study the importance of free secondary schooling for misallocation driven by borrowing constraints in Ghana. Our contribution is to analyze a widespread education policy at the secondary school stage in developed countries: school tracking. In particular, we investigate the consequences of the school track choice and the age at which school tracking occurs for inequality and efficiency in a dynamic macroeconomic model. We thereby complement related research that focuses on the early, pre-school phases in a child's skill development (Daruich, 2022; Yum, 2022) and research that focuses on higher, post-secondary education (Abbott et al., 2019; Capelle, 2022).

Second, this paper builds on the literature on children's skill formation, which has described how children's skills evolve as a function of endowments, parental and environmental inputs, and recently also schooling inputs (see, for instance, Cunha and Heckman (2007); Cunha et al. (2010); Agostinelli et al. (2023, 2019)). Our main innovation relative to this literature is considering two forms of peer effects, which allows for rationalizing the empirical findings regarding school tracking. First, similar to Agostinelli (2018), we incorporate direct peer effects, which capture the idea that children benefit from high-quality peer groups. Second, following Duflo et al. (2011)'s evidence in Kenyan primary schools, we consider how the instruction levels adjust endogenously to the peer composition in schools of a particular track. More specifically, we assume that a child's optimal pace of instruction is unique and increases with her current skill level. Then, learning decreases monotonically with the distance between a child's optimal instruction pace and the one she is currently taught at. This parsimonious micro-funded model captures the main arguments about school tracking and allows us to evaluate the effects of delaying the tracking decision.

Third, this paper contributes to and builds on the literature that estimates the impact of early school tracking on efficiency and equity measures. An extensive empirical literature investigates the effects of age at school tracking on students' test scores and later outcomes. It either exploits temporal within-country variation in tracking practices (Meghir and Palme (2005), for Sweden; Aakvik et al. (2010), for Norway; 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, 2016). Most studies suggest that earlier tracking raises inequality in educational outcomes and increases the effect of parental education on student achievement. Guyon et al. (2012) investigate an educational reform in Northern Ireland that led to a large increase in the share of students admitted to the elite track at age eleven. They find a strong positive overall effect of this de-tracking reform on the number of students passing national examinations at later stages and a negative effect on student performance in non-elite schools that lost their most able students. 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 German students at the margin between two tracks. While their result suggests that the misallocation of hard-to-assign students has little impact on their future outcomes, it does not rule out a potential adverse effect of early school tracking on the outcomes of non-marginal sub-groups of students, such as those from low-socioeconomic backgrounds.

The remainder of the paper is organized as follows. Section 2 presents our life-cycle 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.

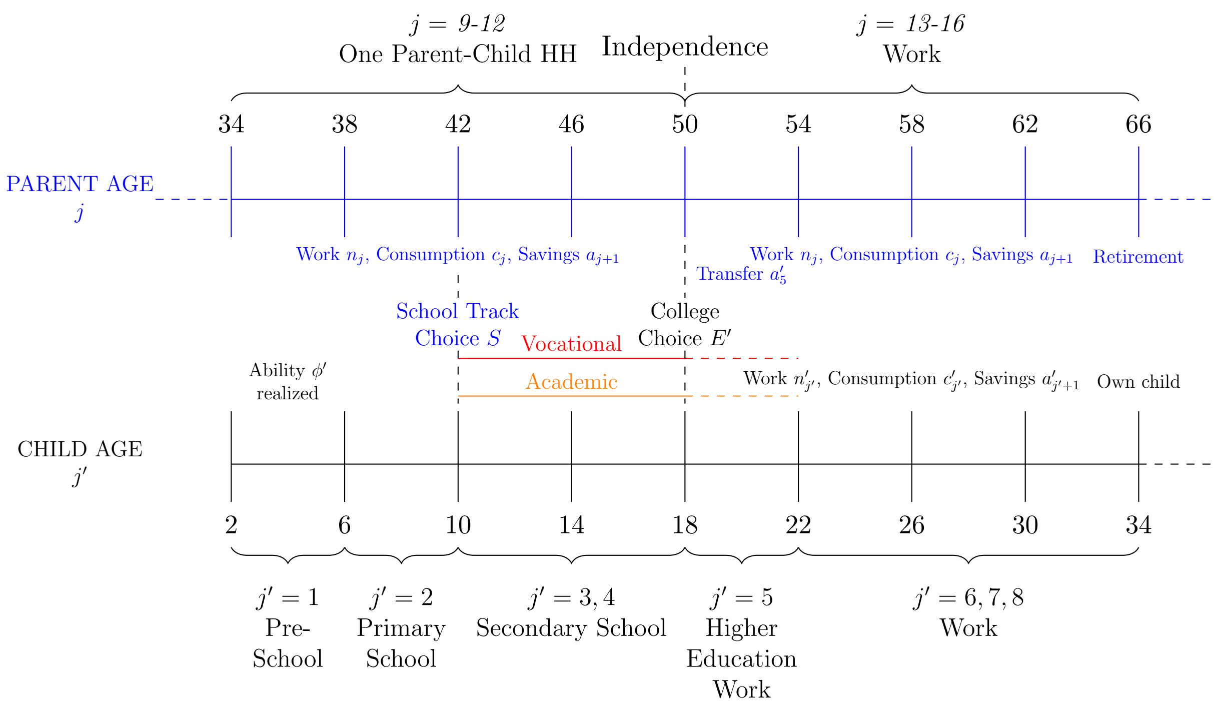
# 2 The Model

Time is discrete and infinite, and one model period, $j \in \{1, ..., 20\}$, corresponds to the 4 years in between ages $[4j-2, 4j+2]$ in real life. Thus, agents enter the model as 2-year-old children and exit at age 82.[5] This frequency allows us to investigate meaningful variations in school tracking ages. The dynastic structure implies that there are 20 generations alive at every point in time. As in Lee and Seshadri (2019), we assume that there is 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 together with her parent, goes to school, and accumulates child skills. School tracking happens in period $j = 3$. In $j = 5$, 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 in the production of the final consumption good. Adult agents decide how much labor to supply until they retire at the beginning of $j = 17$, at age

[5] We choose this perhaps unorthodox timing, such that children are 10 years old when parents make the secondary school track decision, which resembles reality in Germany. An overview of the German Education System is given in Appendix C.

Figure 1: Timeline of Life-cycle Events




66. During the working periods, human capital grows stochastically. Finally, every adult becomes a parent of a new child in $j = 9$, corresponding to age 34, and leaves inter-vivos transfers once her child becomes independent in $j = 13$. We denote the period corresponding to the child by $j'$. Since a generation is 32 years, the child of a parent who is in period $j$ is in period $j' = j - 8$.[6]

## 2.1 Child Skill Formation during the Schooling Years

In period $j = 1$, a new child enters into a one-parent household, equipped with an initial learning ability $\phi'$, which is imperfectly transmitted from her parent (Yum, 2022):[7]

$$\log \phi' = \rho_\phi \ \log \phi \ + \epsilon'_\phi, \qquad \epsilon'_\phi \sim \mathcal{N}(0, \sigma^2_\phi), \tag{1}$$

where $\epsilon_\phi$ is an intergenerational shock. The learning ability translates into an initial

[6]For the remainder of the text, we will denote all child variables with primes.

[7]The learning ability captures genetic components and investments made by parents into their child's skill development during early childhood, infancy, and even in-utero. The importance of these early life stages as well as policy interventions targeted at investments during these years, has been the focus of the child skill development literature (see, e.g., Heckman and Mosso (2014) for a review).

level of (the logarithm of) child skills at the beginning of $j = 2$ when children enter primary school:[8]

$$\theta_2 = \log \phi'. \tag{2}$$

The schooling system during the remainder of the childhood years ($j = 2, 3, 4$) is characterized by the number of distinct school tracks. 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 $S = A$.[9] School tracks differ ex-ante only in their *pace of instruction*, denoted by $P^S$. We think of the pace of instruction as reflecting both different subjects and topics as well as the intensity and depth with which the same topics are taught.[10] The pace of instruction in each track is chosen freely by the policymaker and can be set in order to facilitate her goals. For our analysis, we assume that the goal of the policymaker is to maximize aggregate end-of-school skills.[11] We further assume that there exists a continuum of identical classes (and hence schools) in each track. Thus, if a child is allocated to a particular school track, we can think of her as attending a "representative" class for that track. This implies that all children in a given track are exposed to the same

[8]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. Moreover, we do not allow for potentially different production technologies of cognitive and non-cognitive skills as in Cunha et al. (2010) or Daruich (2022). Instead, in the tradition of Becker and Tomes (1986), we focus on one composite skill, which can be translated into adult human capital that is rewarded on the labor market after school. We do not explicitly model child skill formation during the pre-school stage $j = 1$ as our data on children at this age is sparse.

[9]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. The first age of school tracking varies among OECD countries, from age 10 in Austria and Germany to age 16 in Australia, Canada, Chile, Denmark, Finland, Iceland, New Zealand, Norway, Poland, Spain, Sweden, the United Kingdom, and the United States (OECD, 2013).

[10]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. 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. While we do not necessarily abstract from these factors in affecting child skill development, we conclude that they are not correlated with school track.

[11]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 as well as methods and specific topics that should be taught separately for each school track, subject, and school 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.

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 by:

$$\begin{aligned}&\theta_{j+1} = \kappa\theta_j +\ \alpha\bar{\theta}_j^S +\ g(\theta_j, P_j^S) +\ \zeta E + \eta_{j+1}\\&\eta_{j+1} \sim \mathcal{N}(0, \sigma^2_{\eta_{j+1}}).\end{aligned} \tag{3}$$

Next period's child's skills are directly affected by past skills and parental education, which can be either college or non-college, $E \in \{0, 1\}$. By $\eta_{j+1}$, we denote unobserved i.i.d. shocks to the skills. The existence of this type of uncertainty in the formation of child skills is crucial for the analysis of 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 or even migration shocks that can permanently influence the skill formation trajectory of a child.[12]

Finally, the school track can affect next period's skills in two ways: First, through direct interactions with peers, which is captured through the average skill level of other children in school track $\bar{\theta}_j^S$, similar to Duflo et al. (2011).[13] Second, through the pace of instruction in her school track, governed by the function $g(\theta_j, P_j^S)$. We assume that $g()$ takes the following form:

$$g(\theta_j, P_j) = \beta P_j + \gamma\theta_j P_j - \frac{\delta}{2}P_j^2. \tag{4}$$

Thus, for each skill level $\theta_j$, there exists a unique optimal pace $P^*(\theta_j)$, which, keeping everything else fixed, maximizes future skills. If $\gamma > 0$, this pace strictly increases in current skills, such that higher-skilled children also prefer a higher pace of instruction.

Since there can only be one instruction pace per school track, a policymaker seeking to maximize expected future skills would set the pace in each track to the one that is optimal

[12] 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.

[13] We concentrate on the case with a linear-only direct peer externality governed by $\alpha$. 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 the existence of 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 the network of a child outside of schools (see Agostinelli et al. (2023)), as our data does not contain information on friendships.

for a child with exactly the *average skill level* in that track, as summarized in Lemma 1.

**Lemma 1.** *The pace of instruction in each school track that maximizes expected skills in the next period is given by*

$$P_j^{S^*}(\bar{\theta}_j^S) = \frac{\beta + \gamma \bar{\theta}_j^S}{\delta}, \tag{5}$$

*where $\theta_j^S$ 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 (3) with respect to $P_j^S$ using (4) and under the i.i.d. assumption of $\eta_{j+1}$ and the fact that maximization of skills in each school track is a necessary condition for maximizing unconditional skills. □

Lemma 1 implies that future child skills depend non-linearly on the average skill level of children in her school track. In particular, 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 that she is currently taught at.[14] As a consequence, for a child with a low skill level, going to a school track with a high instruction pace that is tailored to a higher average skill level can be harmful to the point when she actually loses skills despite being surrounded by better peers. The average skill level in a track is not known ex-ante but depends on the distribution of children across tracks.

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

$$h_5 = \exp(\theta_5). \tag{6}$$

## 2.2 Preferences

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

$$u(c_j, n_j) = \frac{(c_j/q)^{1-\sigma}}{1-\sigma} - b\frac{n_j^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}, \tag{7}$$

[14] See Appendix A for the derivation. This formulation is similar to Duflo et al. (2011) and illustrates the efficiency gains in average learning from more homogeneous peer groups. We discuss the theoretical consequences of tracking under these assumptions in detail in Section 3.

where $c_j$ denotes household consumption and $q$ is an adult consumption-equivalent scale that is larger than 1 whenever there is a child in the household and 1 otherwise (Yum, 2022). Risk aversion is captured by $\sigma$. Individuals incur disutility from working hours $n_j$, where $\gamma$ is the Frisch elasticity of labor supply. 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 $\delta$. All future values are discounted by $\beta$.

**Educational Choices**

There are two types of preference shifters related to the two educational choices. First, when deciding on their child's school track, parents' utility is affected by the additively separable constant $\chi(E, S)$, which may depend on their own college education $E$ and their child's school track $S$. The inclusion of this preference shifter is motivated by salient empirical evidence that the school track decision is significantly affected by parental socio-economic status, even conditional on school performance or test scores prior to the track decision, and that parents frequently deviate from the track recommended to them by primary school teachers.[15]

There may be multiple reasons behind these school-track preferences of parents. For example, there may be an information cost associated with acquiring information about the school tracks, and parents may be able to better support their child in a track that they are more familiar with. Parents may also systematically over- or underestimate the potential of their children or have preferences for their child following in their own footsteps. Whatever their exact reason, deviations of parent's track choice from the recommended path may lead to the misallocation of children across tracks. For example, a child with low underlying true potential could be sent to the academic track by parents that 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.

The second education-specific preference shifter concerns the college choice when indi-

[15]See Falk et al. (2020) for evidence in Germany. In our data, we find that children from high-SES parents are 24 percentage points more likely to attend an academic track school, conditional on test scores. Moreover, around 20% of parents who themselves have a college education overrule a recommendation of primary school teachers recommending their child be sent to a vocational track school. See Appendix E for some reduced-form evidence on the school track selection by parental background, deviations from track recommendations, and the consequences of such deviations in terms of later learning outcomes.

viduals reach adulthood ($j = 5$). 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 $\theta_5$ and an idiosyncratic college taste shock, $\nu(E^p) \sim \mathcal{N}(\mu_{\nu,E^p}, \sigma_\nu^2)$, which may be influenced by the parent's education level $E^p$. 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 end up getting a college degree is significantly higher than those with a vocational secondary school degree (so-called "second-chance" opportunities).[16] Secondly, independently of the school track, the likelihood of college education in the data is increasing in the end-of-school skills.[17] Finally, the random taste shocks serve the purpose of reflecting heterogeneity in the higher education decision coming from parental background or from channels that are outside of this model.

## 2.3 Adult Human Capital, Labor Income and Borrowing

During the working career ($j = 5$ to $j = 16$), human capital grows according to (Yum, 2022):

$$h_{j+1} = \gamma_{j,E} \; h_j \; \varepsilon_{j+1}, \quad \log \varepsilon_j \sim \mathcal{N}(0, \sigma_\epsilon^2), \tag{8}$$

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

$$y_j = w_E \; h_j \; n_j, \tag{9}$$

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.

[16] 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 having obtained a higher vocational degree or after having worked for a certain number of years.

[17] Net of the above-explained effect coming through the secondary school track graduation, this may partly be due to the fact that 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 useful.

After retiring, each agent receives retirement benefits $\pi_j(h_{17})$, which depend on the last human capital level before retirement.[18] Finally, the value of death is normalized to zero.

Throughout their life, adult model 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

$$a_{j+1} \geq \frac{-g}{1+r}. \tag{10}$$

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

## 2.4 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, a newly-independent adult decides on her college education in period $j = 5$. A parent decides on the school track of her child in period $j = 11$ and on the inter-vivos transfer in period $j = 13$. All decisions are subject to the human capital growth technology (8), the borrowing constraint (10), a time constraint $n_j \in [0, 1]$ and a period budget constraint, which reads

$$c_j + a_{j+1} = y_j + (1+r)a_j - T(y_j, a_j), \tag{11}$$

where labor income is defined as in (9) and $T(y_j, a_j)$ gives taxes net of transfers, which consist of labor income and capital taxes. We suppress the dependence of the decision problems on these constraints in the following formulations.

### 2.4.1 Parenthood, $j = 9, ..., 13$

**Parent with a young Child ($j = 9,\ 10$)** The state space consists of the parent's education $E$, her human capital, $h_j$, and her assets $a_j$. Moreover, the child's learning ability $\phi'$ is realized in $j = 9$, which corresponds to the first period in the child's life (i.e., $j' = 1$).

[18] 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.

The initial child skill level at the beginning of primary school ($j' = 2$) is given by (2). Future child skills $\theta_{j'+1}$ evolve according to the technology (3) under the assumption of optimal pace setting as defined in Lemma 1. In particular, primary schools are comprehensive track schools, such that the evolution of her skills depends on the average skill level of all children in that cohort $\bar{\theta}_{j'=2}$. The problem of the parent can then be written as:

$$
\begin{aligned}
V_j(E, h_j, a_j, \phi', \theta_{j'}) = \max_{c_j, a_{j+1}, n_j} & \left\{ u(\frac{c_j}{q}, n_j) + \beta\, \mathbb{E}_{\varepsilon_{j+1}, \eta_{j'+1}}\, V_{j+1}(E, h_{j+1}, a_{j+1}, \phi, \theta_{j'+1}) \right\} \\
\text{s.t. } & \theta_{j'+1} = \kappa\theta_j + \ \alpha\bar{\theta}_j^S + \ g(\theta_j, P_j^*(\bar{\theta}_j)) + \ \zeta E + \eta_{j+1}
\end{aligned}
\tag{12}
$$

where expectations are taken over child skill shocks ($\eta_{j'+1}$) and market luck shocks ($\varepsilon_{j+1}$).

**The School Track Decision ($j = 11$)** At the beginning of the period, the parent observes the realization of her child's skills, $\theta_3$, and decides whether to send her child to the vocational or academic track school, $S \in \{V, A\}$. This decision is generally unconstrained by their child's skill, such that even low-skilled children can go to an academic track school.[19] Once a child is tracked, she remains in that track until the end of secondary school (i.e., until $j' = 5$).[20] The track decision is made by comparing the value of sending her child to a vocational track school ($S = V$) with that of sending her child to an academic track school ($S = A$). These values are given by

$$
\begin{aligned}
W_{11}(E, h_{11}, a_{11}, \phi', \theta_3, S) = \max_{c_{11}, a_{12}, n_{11}} & \left\{ u(\frac{c_{11}}{q}, n_{11}) + \beta\, \mathbb{E}_{\varepsilon_{12}, \eta_4}\, V_{12}(E, h_{12}, a_{12}, \phi', \theta_4, S) \right\} \\
\text{s.t. } & \theta_4 = \kappa\theta_3 + \ \alpha\bar{\theta}_3^S + \ g(\theta_3, P_3^{S^*}(\bar{\theta}_3^S)) + \ \zeta E + \eta_4,
\end{aligned}
\tag{13}
$$

for each track $S$. They encode several incentives that influence the track decision. On the

[19] 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.

[20] That is, 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 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. In Dustmann et al. (2017), these "second-chance opportunities" are the main reason why the initial track choice does not have an impact on the marginal child.

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 instruction pace $P_3^{S^*}(\bar{\theta}_3^S)$. As argued above, the latter effect entails a complementarity between own skill and average skills in a school track, which means that children learn better when they are around similar peers.[21] 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 constant shifter $\chi(E, S)$, which describes possibly parent-education-specific preferences for school tracks. Taken together, we can define the value of a parent at the beginning of period $j = 11$ as

$$V_{11}(E, h_{11}, a_{11}, \phi', \theta_3) = \max_{S \in \{V, A\}} \{W_{11}(E, h_{11}, a_{11}, \phi', \theta_3, S) - \chi(E, S)\}, \tag{14}$$

**Remaining Parenthood ($j = 12, 13$)** In period $j = 12$, the parent solves the following problem:

$$\begin{aligned} W_{12}(E, h_{12}, a_{12}, \phi', \theta_4, S) = \max_{c_{12}, a_{13}, n_{12}} & \left\{ u(\frac{c_{12}}{q}, n_{12}) + \beta\, \mathbb{E}_{\varepsilon_{13}, \eta_5}\, V_{13}(E, h_{13}, a_{13}, \phi', \theta_5, S) \right\} \\ \text{s.t. } \theta_5 = \kappa\theta_4 + & \ \alpha\bar{\theta}_4^S + \ g(\theta_4, P_4^{S^*}(\bar{\theta}_4^S)) + \ \zeta E + \eta_5 \ , \end{aligned} \tag{15}$$

where the child's school track $S$ that 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 own market luck shock and her child's final skill shock but does not know the realization of the college taste shock $\nu'(E)$. The transfer cannot be negative, so parents cannot borrow against the future income of their

[21] We show in Section 3 that the track decision of parents who are only interested in maximizing child skills is characterized by a skill threshold, where all children with skills below that threshold go to the vocational track and all children with skills above the threshold go to the academic track.

child. The value at the beginning of period 13 is then

$$\begin{aligned} V_{13}(E, h_{13}, a_{13}, \phi', \theta_5, S) &= \max_{a_5' \geq 0} \left\{ \widetilde{V}_{13}(E, h_{13}, a_{13} - a_5') + \delta \, \mathbb{E}_{\nu'} \, V_{j'=5}(\theta_5, a_5', \phi', S, E) \right\} \\ \nu'(E) &\sim \mathcal{N}(\mu_{\nu,E}, \sigma_\nu^2), \end{aligned} \tag{16}$$

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

$$\begin{aligned} \widetilde{V}_{13}(E, h_{13}, a_{13}) &= \max_{c_{13}, s_{14}, n_{13}} \{ u(c_{13}, n_{13}) + \beta \, \mathbb{E}_{\varepsilon_{14}} \, V_{14}(E, h_{14}, a_{14}) \} \\ \text{s.t. } \; c_{13} + a_{14} + a_5' &= y_{13} + (1+r) a_{13} - T(y_{13}, a_{13}), \end{aligned} \tag{17}$$

so that the transfer $a_5'$ enters the budget constraint.

### 2.4.2 Work Life Without a Dependent Child, $j = 5, 6, 7, 8$ and $j = 14, 15, 16$

**Independence ($j = 5$)** At the beginning of adulthood ($j = 5$), the state space of a newly independent adult comprises the secondary school track she graduated from $S$, end-of-school skills $\theta_5$, initial assets $a_5$, which she received from her parents, innate ability $\phi$ and her parent's education $E^p$, which affects the mean 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:

$$\begin{aligned} V_5(\theta_5, a_5, \phi, S, E^p) = \max\{ & W_5(E = 0, h_5, a_5, \phi), \\ & W_5(E = 1, h_5, a_5, \phi) - \psi(S, \theta_5, \nu(E^p)) \} \end{aligned} \tag{18}$$

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

$$W_5(E, h_5, a_5, \phi) = \max_{\substack{c_5, a_6 \\ n_5 \in [0, \bar{n}(E)]}} \{ u(c_5, n_5) + \beta \, \mathbb{E}_{\varepsilon_6} \, V_6(E, h_6, a_6, \phi) \} \tag{19}$$

and end-of-school skills are transformed into adult human capital $h_5$ according to (6). 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** $j = 14, 15, 16)$ For the two following periods $j = 6, 7$, an adult without a child solves the following life-cycle savings problem:

$$V_j(E, h_j, a_j, \phi) = \max_{c_j, a_{j+1}, n_j} \left\{ u(c_j, n_j) + \beta \, \mathbb{E}_{\varepsilon_{j+1}} V_{j+1}(E, h_{j+1}, a_{j+1}, \phi) \right\}. \tag{20}$$

In period $j = 8$, the individuals know that they will become parents next period. For that reason, they take expectations over the learning ability of their future child, $\phi'$, on top of the expectations over the market luck shocks. Ability is imperfectly transmitted from parents to children, according to (1). Thus, the value in period 8 becomes

$$V_8(E, h_8, a_8, \phi) = \max_{c_8, a_9, n_9} \left\{ u(c_8, n_8) + \beta \, \mathbb{E}_{\varepsilon_9, \phi' | \phi} V_9(E, h_9, a_9, \phi') \right\} \tag{21}$$

For the remainder of the periods $j = 14, 15, 16$, an adult whose child has left the household solves the standard life-cycle savings problem:

$$V_j(E, h_j, a_j) = \max_{c_j, a_{j+1}, n_j} \left\{ u(c_j, n_j) + \beta \, \mathbb{E}_{\varepsilon_{j+1}} V_{j+1}(E, h_{j+1}, a_{j+1}) \right\}, \tag{22}$$

where the learning ability $\phi$ has already been transmitted to the child and does not enter the state space anymore.

### 2.4.3 Retirement, $j = 17, 18, 19, 20$

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

$$\begin{aligned} V_j(E, h_{17}, a_j) &= \max_{c_j > 0, a_{j+1} \geq \underline{a}} \{ u(c_j, 0) + \beta V(E, h_{17}, a_{j+1}) \} \\ \text{s.t. } c_j + a_{j+1} &= \pi_j(h_{17}) + (1 + r) a_j - T(0, a_j). \end{aligned} \tag{23}$$

## 2.5 Aggregate Production, and Government

We assume that 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 a CES aggregate of total labor supply, which is defined by:

$$H = [\varphi H_0^{\epsilon} + (1 - \varphi) H_1^{\epsilon}]^{\frac{1}{\epsilon}}. \tag{24}$$

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 $\delta_f$.

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 $\tau_a r a_j$ (Yum, 2022). All tax revenue is used to finance retirement benefits $\pi_j$ as well as fixed lump-sum social welfare benefits $g$ that are paid to every household. These may include child allowances, unemployment benefits, or contributions to health insurance.

## 2.6 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 linearly independent spending.

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 average skill levels in each track, which in equilibrium, coincide with the actual distributions. A detailed definition of the equilibrium is given in Appendix B.

# 3 Developing Intuition: School Tracking and Child Skill Formation

Our formulation of the child skill formation technology during the schooling years in (3) and (4) constitutes the novel cornerstone of our model. In order to develop some intuition about the main mechanisms surrounding school tracking and child skill formation at work in our model, we use this section to derive a series of theoretical implications about the effects of school tracking policies on the distribution of end-of-school skills, that follow from the child skill technology. Consequently, 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, $\zeta = 0$. Thus, skills evolve according to

$$\begin{aligned} &\theta_{j+1} = \theta_j + \alpha\bar{\theta}_j^S + \beta P_j^S + \gamma\theta_j P_j^S - \frac{\delta}{2}(P_j^S)^2 + \eta_{j+1} \\ &\eta_{j+1} \sim \mathcal{N}(0, \sigma^2_{\eta_{j+1}}). \end{aligned} \tag{25}$$

All other assumptions are maintained. In particular, policymakers set the instruction paces in each school track with the goal of maximizing 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.[22]

## 3.1 Comprehensive School versus Tracking

We start by comparing a comprehensive schooling system, in which all children attend the same school track, with a tracking system in which all children are tracked into a vocational or academic track. For simplicity, we consider only one period of schooling here. Thus, if $\theta_3$ are the skills at the beginning of secondary school, $\theta_4$ can be thought of as the skills at the end of school. We begin by describing how the allocation of children across tracks in the tracking system happens optimally.

### 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 ($\mathbb{E}(\theta_4)$).

The second alternative consists of each parent making the track decision unilaterally for her child, which is characterized by a skill level $\theta_{i,3}$. A parent's only goal is to maximize her child's expected end-of-school skill level ($\mathbb{E}(\theta_{i,4})$). Parents know the distribution of $\theta_3$.

[22]This seems a good approximation as the achievement test scores that we use as measures for child skills are indeed bell-shaped and centered around zero in our data.

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 payoffs are given by the next period's skills.

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.[23] Intuitively, this generates the highest aggregate end-of-school skills as it minimizes the variance of skills in each track, or in other words, 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 exactly 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 make their decision irrespective 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 is exactly equally well off in both tracks. The location of this threshold is smaller than the optimal threshold a policymaker would pick whenever the direct peer effects are positive ($\alpha > 0$). Intuitively, parents do not internalize the negative effect that this deviation from the optimal threshold has on aggregate end-of-school skills.

**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 track allocation is done by the policymaker, the optimal skill threshold corresponds to the average initial skill level* $\tilde{\theta}_3^* = \mathbb{E}[\theta_3] = 0$.
- *If the track allocation is done by parents, the endogenous skill threshold that emerges from this game depends on the direct peer externality* $\alpha$*. With* $\alpha > 0$*, the threshold is smaller than* $\tilde{\theta}_3^*$*.*[24]

*Proof.* In Appendix A. □

Next, we compare the comprehensive and tracking system in terms of their effects on

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

[24]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.

end-of-school skills, assuming that tracking happens optimally, as described 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 peer groups in each track in terms of their initial skills. Since learning decreases in the variance of skills among children in a track, more homogeneity on average increases end-of-school skills.

The advantage of tracking in terms of increasing aggregate end-of-school skills increases in the complementarity between own skills and instruction pace, $\gamma$. 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 $\delta$, 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 (26) 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, a full tracking system necessarily leaves a non-negative mass of 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.[25] 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 especially the children at the bottom of the skill distribution suffer from a tracking system (see, 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*

[25] This is interesting in a political economy context as the median voter in this model would prefer a comprehensive system. This could partially explain why we see different tracking systems across different countries.

*of the tracking skill threshold $\tilde{\theta}_3$.*

- *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*

$$\alpha^2 + 2\alpha\left(1 + \frac{\beta\gamma}{\delta}\right) - (8-\pi)\frac{\gamma^4}{\pi\delta^2}\sigma^2_{\theta_3} > 0. \tag{26}$$

- *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. □

Note that these results are not affected by the presence of skill shocks in the one-period model. This is because these shocks are assumed to be mean zero and realized at the end of the period. The role of uncertainty changes once tracking remains for more than one period.

## 3.2 Early versus Late Tracking

Let us now consider a two-period secondary schooling system so that $\theta_5$ are the skills at the end of school, exactly like in our full model. 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 which the allocation is done optimally by the policymaker. The early tracking system is characterized by an initial track allocation into two tracks, $V$ and $A$, at the beginning of secondary school, $j = 3$. In an early tracking system, a child therefore remains in her school track for the two secondary school periods. The late tracking system is characterized by all children going to a comprehensive school in the first period, followed by tracking into $V$- or $A$-track schools at the beginning of the second secondary school period ($j = 4$). Importantly, while this allocation occurs after the skill shock $\eta_4$ is realized in the $LT$ case, in the $ET$ case, the allocation occurs before.[26]

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.

[26] We do not consider a fully comprehensive system in which children remain in comprehensive track schools for the whole duration of their school career. Proposition 2 implies that such a system cannot achieve higher aggregate end-of-school skills compared to a late tracking system.

Intuitively, this represents the key disadvantage of early tracking. Since the first track selection cannot be corrected for unexpected skill shocks, the peer groups in each track become more heterogeneous. Consequently, the average skill levels and hence the instruction paces across tracks are closer together than what would be optimal if re-tracking was possible. Thus, while early tracking produces learning gains in the first stage, it takes away the flexibility to react to unexpected changes in the composition of the tracks.

**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*

$$\frac{\sigma^2_{\eta_4}}{\sigma^2_{\theta_3}} > 1 + \alpha + \alpha^2 + \beta + \frac{\beta^2}{2} + 2\alpha(1+\beta) + \frac{\gamma^2}{2\pi}\sigma^2_{\theta_3}. \tag{27}$$

*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, 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.

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 preferences about 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 parameterize the model 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. A summary of the externally calibrated parameters is given in Table 3 and of the internally estimated ones in Table 4.

## 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 children in nationally representative samples in Germany (Blossfeld et al., 2011). A key component of the information collected is regular standardized assessment tests of the children's competencies in areas such as mathematics, reading, sciences, vocabulary, or grammar, combined with specific wave weights. In addition, there is information about school track recommendations and the final parental school track choices.[27] We further restrict the sample to individual observations containing information on the school and class in that school a child attended in a given year.

The second data source is the German Socio-Economic Panel (SOEP), an annual representative survey from which we use the 2010-2018 waves. 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 that we do is dropping those with hourly wages below the first and above the 99th percentile. Lastly, 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 important ingredient of our model. Then, we describe the functional forms and estimation strategies for all remaining parameters.

[27]See also Appendix Section D for more details on the tests.

## 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:[28]

$$\theta_{i,j+1} = \omega_{0,j} + \omega_{1,j}\theta_{i,j} + \omega_{2,j}\bar{\theta}^S_{-i,j} + \omega_{3,j}{\theta_{i,j}}^2 + \omega_{4,j}(\theta_{i,j} - \bar{\theta}^S_j)^2 + \omega_{5,j}E_i + \eta_{i,j+1}, \tag{28}$$

Note that (28) is just a rearranged version of the child skill technology (3) after substituting in (4) and the optimal pace of instruction in each school track as given by Lemma 1 and after adding and subtracting $\frac{\gamma^2}{2\delta}\theta^2_{i,j}$.[29] Moreover, we, in principle, allow all child skill technology estimates to be specific to the period $j$.[30]

We estimate this version of the technology because it depends solely on observable information in the data, and not on the "constructed" instruction paces across school tracks. Moreover, since 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, we cannot identify a model that includes $\theta\bar{\theta}_j$, $(\bar{\theta}^S_j)^2$, and the interaction $\theta\bar{\theta}^S_j$, as in (3). Instead, we rewrite (3), such that it depends on the squared differences between own skills and average skills in a school track. This restricts the coefficients $\omega_3 = -\omega_4$, which we formally test after the estimation.

In the estimation, we also distinguish between $\bar{\theta}^S_{-i,j}$, which denotes the average skill level of 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, $\bar{\theta}_{-i,j,S} = \bar{\theta}_{j,S}$, since we assume a representative school and class per track (or alternatively, identical classes conditional on school tracks). In the data, however, there is clearly heterogeneity across classes, even within a school track. 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. Finally, the intercept $\omega_{0,j}$ can be a function of age and gender in the empirical estimation, and the parental educational attainment $E$

[28]Following the work in Cunha et al. (2010), much of the empirical and quantitative literature using child skill formation technologies employed 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.

[29]The coefficients $\omega_{n,j}$, $n = 0, ..., 5$ relate to those in (3) and (4) as follows: $\omega_0 = \frac{\beta^2}{2\delta}, \omega_1 = (\kappa + \frac{\beta}{\gamma}\delta), \omega_2 = \alpha, \omega_3 = -\omega_4$, and $\omega_5 = \zeta$ for all $j$.

[30]*Note that in the current version of the paper, we have used only one cohort of the NEPS data to estimate the technology. We use the estimates from that cohort for all periods.*

is a time-constant dummy that equals 1 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 $\theta$ as latent variables that are only imperfectly measured in the data. For that reason, we employ a log-linear measurement system for latent skills, using a series of achievement test scores as noisy measures of child skills in each period.[31] The identification strategy of the scales and loadings of each measure using their covariances follows Cunha et al. (2010). To account for measurement error, we aggregate the individual measure into a composite unbiased index using Bartlett factor scores, as in Agostinelli et al. (2023). Appendix F provides details on skills measurement and the estimation procedure.

We present our preferred estimates of the child skill production technology parameters in Table 1. Note that the most reliable estimates are based on the data from NEPS Starting Cohort 3, which follows children through secondary school. For that reason, we assume for now that the estimates of the child skill technology parameters between school grades 5 and 9 are representative of the entire schooling career. That is, we drop the $j$ index on the technology parameters. Recall that $\theta_{i,j}$ is defined as the logarithm of child skills. Hence, we can interpret the coefficients as elasticities. Thus, $\hat{\omega}_1 = 0.65$ means that a 1% increase in latent skills at the beginning of primary school is associated with an 0.65% increase with end-of-primary school skills. This own-skill productivity is close to values commonly found in the literature (see estimates in Cunha et al. (2010); Agostinelli et al. (2019)). During secondary school, the estimated coefficient $\hat{\omega}_2$ is positive. More importantly, we cannot reject the hypothesis that $\hat{\omega}_2 = -\hat{\omega}_4$ which is in line with our assumptions in Section 3.

The estimated coefficient $\hat{\omega}_4$ is negative and statistically significant at the 10% level. It indicates that a 1% increase in the squared distance to the average skill level in a track is associated with an up to 0.05% 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, in both directions, from this level can hurt individual skill development. The estimated $\hat{\omega}_3$ are generally small and often statistically insignificant.

The final estimates we use to parameterize the child skill formation technology in our model are then $\omega_n$ for $n = 1, 2, 4, 5$ as reported in Table 1. The intercept is calibrated

[31] 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.

internally, such that average log skills are always zero in the baseline model, which is one of our identifying assumptions.

Table 1: Child Skill Technology Parameters Estimates

| Dependent Variable: $\theta_{i,j+1}$ Grade 9 on Grade 5 | | |
|---|---|---|
| Coefficient | Variable | |
| $\hat{\omega}_1$ | $\theta_{i,j}$ | 0.65 (0.026) |
| $\hat{\omega}_2$ | $\theta_{i,j}^2$ | 0.02 (0.02) |
| $\hat{\omega}_3$ | $\bar{\theta}_{-i,j,S}$ | 0.12 (0.082) |
| $\hat{\omega}_4$ | $(\theta_{i,j} - \bar{\theta}_{j,S})^2$ | -0.05 (0.025) |
| $\hat{\omega}_5$ | $E = 1$ | 0.10 (0.04) |
| $N$ Children | | 1,675 |

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

## 4.3 Remaining Parameters

### Preferences

We set the inverse elasticity of intertemporal substitution, $\sigma = 2$, 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 in order to match the average time worked in the SOEP data, given that the total time available after sleep and self-care is normalized to 1.

We set the time discount factor $\beta$, such that the equilibrium interest rate amounts to 4% annually. The altruism parameter $\delta$ is calibrated such that the ratio of average inter-vivos transfers to average labor income in the model corresponds to that of average higher education costs of children to average 4-year labor income in the data. According to a survey by the German Student Association in 2016, the monthly costs of living during the higher education stages range from 596 to 1250 Euros (Middendorf 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 5 years (the length of time it takes to complete studies that are equivalent to a master level). Then, the ratio of total costs to average 4-year labor income ranges from 0.32 to 0.67. In our baseline calibration, we take as a target a ratio of 0.6.

We parameterize the school track choice preference shifters $\chi(E, S)$ as follows

$$\chi(E,S) = \begin{cases} \chi_1 & \text{if } E = 1 \wedge S = A \\ \chi_0 & \text{if } E = 0 \wedge S = V, \end{cases} \tag{29}$$

so that $\chi_1$ represents a preference for the academic track of college-educated parents, and $\chi_0$ represents a preference for the vocational track of non-college-educated parents. We estimate these parameters to match the share of deviations from secondary school track recommendations by parental education in the data. These recommendations are typically given by primary school teachers before the transition to secondary school. They are based on both a reflection of the child's achievement during primary school as well as 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 the primary school teachers typically observe the children over multiple years every day during the week, based on a similar information set as the parents possess. For that reason, we think of the recommended school track in the model as the one which a parent would have chosen with $\chi_1 = \chi_0 = 0$. Then, deviations from that unbiased track choice by parental education map into deviations from teacher recommendation.

**Initial Child Skills and Child Skill Shocks**

Initial child skills just before entering primary school are a function of the learning ability of a child, which is imperfectly transmitted from the parent following an AR(1) process with inter-generational correlation coefficient $\rho_\phi$, and variance $\sigma^2_\phi$. Since the learning ability is correlated with the eventual higher education outcome of a parent, we pick as the target moment for $\rho_\phi$ the difference in average preschool child skills by parental education (measured in one standard deviation). The variance $\sigma^2_\phi$ is then estimated to match the variance of initial math test scores in the data.

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 math test score 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.[32]

### 4.3.1 College Costs

We paramterize the "psychic" college cost function as follows (Daruich, 2022):

$$\begin{aligned} \psi(S, \theta_5, \nu(E^p)) &= \exp(\psi_0 + \psi_{S=V} + \psi_\theta \theta_5 + \nu(E^p)) \\ \nu(E^p) &\sim \mathcal{N}(\mu_{\nu,E^p}, \sigma^2_\nu). \end{aligned} \tag{30}$$

We estimate the two parameters $\psi_0$ 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 $\psi_\theta$ that multiplies end-of-school skills by matching the regression coefficient on log math test scores from a regression of a college graduation dummy on end-of-school test scores.

We calibrate the two parental education-specific means of the college taste shock parameters, $\mu_{\nu,E^p=0}$ and $\mu_{\nu,E^p=1}$, to match the share of children from each parental education background that receive a college degree in the data. Finally, we calibrate the variance of these shocks, $\sigma^2_\nu$, to match the variance of the residuals from the above regression of college education on end-of-school skills, as in Daruich (2022).

The final component of college costs is not a part of the "psychic" costs $\psi$ 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.[33] Thus, we set the maximum remaining time during the higher education stage to

[32] In reality, such changes may also arise from factors that are outside the scope of this model but can put children on a different skill formation path. These could be, for example, a change of schools within a school track, a change of teachers within a class, or even tutoring sessions that are uncorrelated with parental education.

[33] A common 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 8 semesters or 4 years.

$\bar{n}(E=1) = 0.40$.

### 4.3.2 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).[34] The resulting experience-wage profiles for 4-year experience bins are shown in Table 2, expressed in growth relative to the previous bin. We set the $\{\gamma_{j,E}\}_{j=5}^{16}$ parameters to these values.

Finally, we calibrate the variance of the market luck shocks, $\sigma_\varepsilon^2$, such that our model replicates the standard deviation of labor income across all workers in the data.

Table 2: 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

[34] Concretely we create, separately for each education group, 4-year work experience bins. We then estimate Mincer regressions of wages on years of schooling and potential work experience, controlling for 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

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

To disentangle time from cohort effects, we assume that there is no experience effect on wage growth in the last 8 years of work, following the HLT approach in Lagakos et al. (2018).

### 4.3.3 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 vocational and academic human capital in the firm production is equal to 1.5 (Ciccone and Peri, 2005). The weight on vocational human capital in the CES aggregator, $\varphi$ 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 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 in the economy. Finally, we set pension benefits to $\pi_j = \Omega h_j w_E$ during retirement and calibrate the scale parameter $\Omega$ internally, such that the average replacement rate corresponds to 40%.

Table 3: Parameters calibrated externally

| Parameter | Value | Description | Source |
|---|---|---|---|
| **Household** | | | |
| $\sigma$ | 2.0 | Inverse EIS | Lee and Seshadri (2019) |
| $\gamma$ | 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 | |
| **Firm** | | | |
| $\sigma_f$ | 1/3 | E.o.S Vocational and Academic Human Capital | Ciccone and Peri (2005) |
| $\delta_f$ | 6% | Annual Depreciation | Kindermann et al. (2020) |
| **Government** | | | |
| $\tau_n$ | 0.128 | Labor Tax Progressivity | Kindermann et al. (2020) |
| $\lambda$ | 0.679 | Labor Tax Scale | Kindermann et al. (2020) |
| $\tau_a$ | 0.35 | Capital Tax Rate | |
| $g$ | 0.06 | Lump-sum Transfers | |

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

## 4.4 Method of Simulated Moments Estimation Results

In total, we estimate 20 parameters internally using the method of simulated moments to match 20 target data moments. The parameters, their estimated values, model-implied moments, and target data moments are presented in Table 4.

Table 4: Internally Calibrated Parameters

| Parameter | Value | Description | Target | Data | Model |
|---|---|---|---|---|---|
| **Preferences** | | | | | |
| $\beta$ | 0.935 | Discount Factor | Annl. Interest Rate | 0.04 | 0.04 |
| $b$ | 6.8 | Labor Disutility | Avrg. Labor Supply | 0.53 | 0.53 |
| $\delta$ | 0475 | Parental Altruism | Transfer/Income | 0.60 | 0.59 |
| $\chi_0$ | 0.017 | Own V-Track Bias | Share of Deviations | 0.16 | 0.18 |
| $\chi_1$ | 0.021 | Own A-Track Bias | Share of Deviations | 0.23 | 0.21 |
| **College Costs** | | | | | |
| $\psi$ | 0.88 | Intercept | Share A $\rightarrow$ College | 0.71 | 0.71 |
| $\psi_V$ | 0.25 | Add. Costs for V-Track | Share V $\rightarrow$ College | 0.11 | 0.08 |
| $\psi_\theta$ | 0.74 | Coefficient on $\theta_5$ | Regression Coefficient | 0.80 | 0.94 |
| $\mu_{E^p=0}$ | 0.1 | Mean Taste Shock if $E^p = 0$ | Share in CL from Non-CL HH | 0.20 | 0.18 |
| $\mu_{\nu,E^p=1}$ | -0.1 | Mean Taste Shock if $E^p = 1$ | Share in CL from CL HH | 0.64 | 0.66 |
| $\sigma_\nu$ | 0.001 | Std. Taste Shock | Variance of Residual | 0.218 | 0.122 |
| **Idiosyncratic Shocks** | | | | | |
| $\sigma_\varepsilon$ | 0.008 | Std. Luck Shock | Std(Log Labor Income) | 0.73 | 0.82 |
| $\sigma_\phi$ | 0.07 | Std. Ability Shock | Var(Test Scores Grade 1) | 0.12 | 0.12 |
| $\rho_\phi$ | 0.65 | Persistence of Ability | Test Score Diff. (Grade 1) by $E$ | 0.47 | 0.52 |
| $\sigma_{\eta_3}$ | 0.07 | Std. Learning Shock $j = 3$ | $\text{Rank}_{j=2}$-$\text{Rank}_{j=3}$ | 0.59 | 0.60 |
| $\sigma_{\eta_4}$ | 0.06 | Std. Learning Shock $j = 4$ | $\text{Rank}_{j=3}$-$\text{Rank}_{j=4}$ | 0.63 | 0.68 |
| $\sigma_{\eta_5}$ | 0.05 | Std. Learning Shock $j = 5$ | $\text{Rank}_{j=4}$-$\text{Rank}_{j=5}$ | 0.72 | 0.74 |
| **Miscellaneous** | | | | | |
| $\Omega$ | 0.14 | Pension Anchor | Replacement Rate | 0.40 | 0.39 |
| $A$ | 2.5 | TFP | Avrg. Labor Earnings | 1.0 | 1.0 |
| $\omega$ | 0.54 | 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.

The model generally 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 34.7%, which is in line with the German data in the 2010s. The model also matches well the transition rates from academic and vocational secondary school into college higher education (at around 70% and 10%), which implies that the share of children in an academic track school in the model, 42%, is in accordance with the data.

Parental preferences towards their own track affect the school track decision significantly, both in the model and in the data. In particular, around 20% of parents from each education background overrule a different track recommendation by teachers in the NEPS data. In the model simulated data, roughly the same shares of parents would send their child to a different track if it was not for the preference shifter $\chi(E, S)$. Furthermore, the correlation between parental and child college education in the model matches the data.

In order to match the correlation between child skill ranks across school periods, the model requires rather large child skill shocks. This is in part because the estimated own-skill productivity, $\omega_1$, in the child skill formation technology is also quite large. The model resembles well the differences in initial child skills by parental education prior to entering school. In particular, while children from college-educated parents have an average initial skill level that is around 0.47 standard deviations larger than the average level of non-college-educated parents in the data, this difference is 0.52 standard deviations in the model.

## 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 5. The first set of moments pertains to child skills. While we target the difference (in terms of standard deviations) in average initial child skills prior to entering primary school in the calibration, we do not track how this difference evolves over the school career. In both data and model, the differences in parental education and school track increase slightly during secondary school.[35]

Similarly, the differences in average child skills across school tracks (in terms of standard deviations) increase slightly throughout secondary school in the data. In our model, however, these differences, while remaining large, slightly decrease over time. In general, differences across school tracks are generally larger than differences across parental education.

The second set of moments concerns the relationship between track choice and parental education. In the data, the share of college-educated parents who send their child to an

[35] 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.

academic track school is around 74%, which is similar to what our model produces. For non-college-educated parents, this share is only 24%. Moreover, we regress a dummy variable that equals one if a child attends an academic track school on the percentile rank of the child's skills prior to secondary school in order to assess the skill gradient in academic track choice. The estimated coefficient is 0.87 in the data and 1.02 in the model, suggesting that our model slightly overestimates the importance of child skills for the track choice.

The third set of moments relates to intergenerational mobility. 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 10 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 5.0 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, which our model matches well. We also compare our model-implied estimate of the intergenerational elasticity of income (IGE) to estimates on German data by Kyzyma and Groh-Samberg (2018). Compared to their findings, the model produces IGEs that are within the range of their data counterparts.

Finally, the model understates the degree of inequality in labor incomes as measured by the Gini coefficient slightly. The ratio of average earnings among college-educated to non-college-educated workers is consistent with the data.

**Long-term effects of Track Choice for Marginal Students**

Dustmann et al. (2017) analyse 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 that are born just before the cut-off age are less likely to go to an academic track secondary school, simply because they are younger at the time of the track decision relative to their class peers. 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.[36]

[36] 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

Table 5: Non-targeted moments

| Moment | Data | Model |
|---|---|---|
| **Child Skill Moments** | | |
| Mean Differences by Parental Background (in Standard Deviations) | | |
| Beginning Secondary School | 0.58 | 0.65 |
| Middle Secondary School | 0.70 | 0.71 |
| Mean Differences by School Track (in Standard Deviations) | | |
| Beginning Secondary School | 0.87 | 0.92 |
| Middle Secondary School | 1.01 | 0.77 |
| **School Track Choice** | | |
| Relative share A-track children from CL. HH | 0.74 | 0.72 |
| Relative share A-track children from Non-CL HH | 0.24 | 0.25 |
| Coefficient A-track on Skill Rank | 0.87 | 1.02 |
| **Intergenerational Mobility** | | |
| Parental Income Gradient (Dodin et al., 2021) | 0.52 | 0.50 |
| Q5/Q1 A-track on income (Dodin et al., 2021) | 2.13 | 2.50 |
| Q1 A-track on income (Dodin et al., 2021) | 0.34 | 0.26 |
| IGE (Kyzyma and Groh-Samberg, 2018) | 0.27-0.37 | 0.30-0.33 |
| **Inequality - Returns to College** | | |
| Gini Coefficient of Labor Income | 0.29 | 0.26 |
| CL/Non-CL Earnings | 1.69 | 1.76 |

*Notes*: This table presents the non-targeted moments and their model-generated counterfactuals.

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 that are very similar in terms of their state variables at the point of school track choice but end up going to different school tracks. Naturally, in our model, we cannot distinguish the date of birth for children of the same cohort. For that reason, we distinguish children by their skills prior to the secondary school track choice ($\theta_{j=3}$). As detailed in Section 3, our child skill development technology implies that, conditional on parental background, the school track choice is characterized by a skill threshold, such that all children with skills above that threshold go to the academic school track and all below go the vocational track school. Conditional on all other states at the time of the track choice – parental human capital, assets, education, and learning ability – differences in child skills and hence differences in school track choice in our model arise from randomly drawn skill shocks. Analogously to Dustmann et al. (2017), we could alternatively argue that these shocks are (at least partly) the result of within-cohort age differences of children, which affect their skill development but are not explicitly modeled. Thus, comparing the later-in-life outcomes of otherwise very similar children with skills

will have accumulated less work experience.

around the tracking threshold can be interpreted as estimating the effect of school track choice induced by random (age or skill) shocks.

Concretely, we compare children with skills in a 10% interval around the tracking threshold who go to different school tracks, conditional on all other states.[37] We evaluate these marginal children in terms of their labor income at age 30, the present value of their lifetime labor income, and the present value of their lifetime wealth.[38] We find that going to the academic track instead of the vocational track is associated with a 10.4% higher labor income at age 30, a 4.5% higher present value of lifetime labor income, and a 4.6% higher present value of lifetime wealth.

While not zero, these differences seem rather small in relation to overall inequality in these outcomes. For example, the 4.5% higher present value of lifetime labor income is around 1/20th 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 that children at the margin of these two vocational tracks 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

The benefit of our model is that we can use it 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 sources of lifetime and inter-generational inequality in the spirit of Huggett et al. (2011) and Lee and Seshadri (2019). Then, we investigate the determinants and consequences of secondary school track choice, as this constitutes the main novelty of our model. In this context, we perform counterfactual analysis of economies in which the school track decision is not affected by parental preferences or in which a policymaker enforces a

[37] This interval amounts to around 1/5 of a standard deviation of child skills prior to the school track choice. We form quintiles of the continuous states of parental human capital and parental assets and allocate children into discrete groups pertaining to these quintiles. Moreover, we partition the distribution of the learning ability $\phi'$ into three ability states. For these reasons, the skill threshold can become fuzzy in the sense that even conditional on these groups, a child with slightly higher skills goes to the vocational track, whereas a child with slightly lower skills goes to the academic track.

[38] 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.

strict tracking skill threshold. Finally, we study the effects of a counterfactual policy reform that postpones the school tracking age to 14.

## 5.1 Sources of Inequality

Using our model, 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. Following the literature, we focus on lifetime labor income and lifetime 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 lifetime labor income and wealth.[39] These states are the school track in secondary school $S$, initial adult human capital $h_5$, initial transfers received from the parent $a_5$, the child's college choice $E$, parental college education $E^p$, and innate learning ability $\phi$.

Table 6: Contributions to Lifetime Inequality

| | | Share of Explained Variance | |
|---|---|---|---|
| Life Stage | States | Lifetime Earnings | Lifetime Wealth |
| Independence (age 18) | $(S, \phi, h_5, a_5, E, E^p)$ | 70% | 65% |
| | $(S, \phi, h_5)$ | 63% | 60% |
| | $(S, \phi, a_5, E, E^p)$ | 54% | 45% |
| School Track Choice (age 10) | $(S, \phi', \theta_3, h_{11}, a_{11}, E)$ | 23% | 30% |
| | $(S, \theta_3, \phi')$ | 20% | 21% |
| | $(S)$ | 12% | 13% |
| Pre-Birth (parent age 30) | $(E, \phi, h_8, a_8)$ | 10% | 20% |

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

Row 1 of Table 6 summarizes that 70% 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%.[40] Thus, our model suggests that the majority of lifetime outcomes is already predetermined when agents become independent and can enter the labor market. Note that at this stage, all uncertainty regarding initial human capital as well as the college decision has been made. The remaining unresolved uncertainty over human capital (market luck) shocks during the working years has, therefore, only limited effects on lifetime inequality.

[39]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. As before, we partition the continuous states into three equally sized groups.

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

As Row 2 of Table 6 shows, 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: the child's secondary school track $S$, her learning ability $\phi$ and her end-of-school skills that are transformed into initial adult human capital, $h_5$. This suggests that the size of the parental transfer $a_5$ and the college choice $E$, even when affected by parental education $E^p$ are not major sources of lifetime inequality. Instead, if we only exclude initial adult human capital $h_5$ (Row 3), the share of explained variance in lifetime earnings drops by almost 16 percentage points, and the share of explained variance in lifetime wealth by 20 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.[41] Interestingly, the correlation between initial adult human capital and transfers received from parents is negative in the model. This suggests that parents partially offset the disadvantage their children experience in the labor market from having lower skills by giving them higher transfers.[42]

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 23% of lifetime earnings and 30% of lifetime wealth variation is explained (Row 4). Again, the majority of this variation seems attributable to differences in child states at that age. Yet the explained share is clearly 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 12% of lifetime earnings variation and 13% 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. In fact, as we argued in Section 4.5, for children with similar skills, the track choice has only small independent effects on lifetime outcomes. We investigate the determinants and consequences of the school track choice in more detail below.

The last row of Table 6 shows the contribution of parental states prior to the birth of their children to their children's lifetime outcomes. At this stage, none of the uncertainty regarding child skill and human capital shocks nor regarding the child's learning ability has been realized. Still, around 10% 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

[41] We cannot, however, attribute these drops exclusively to initial human capital differences, given the possible correlation between states.

[42] This channel is also present, albeit to a smaller degree, in Lee and Seshadri (2019).

wealth, this share is even higher at 20%, pointing to the important role of wealth transfers. For example, using the same decomposition of the unconditional variance of transfers into parental states pre-birth, we find that almost 31% of variation in transfers is predetermined prior to the birth of the child. In contrast, only 16% of the variation in human capital at age 18 is predetermined prior to birth, which highlights the role of shocks to child skills during their childhood and school years.

## 5.2 School Track Allocation Counterfactuals

According to the theoretical predictions laid out in Section 3, the initial school track should, to a large degree, be based on child skills. A regression of an academic school track dummy on all states at the time of the tracking decision using model-generated data confirms that this is true in our model. Column 1 of Table 7 reports the standardized coefficient estimates of this regression, indicating that child skills at the time of the track choice, $\theta_3$ have the strongest impact on the track decision. In particular, increasing log child skills by one standard deviation increases the probability of going to the academic track by 53 percentage points.

Notwithstanding this, Column 1 in Table 7 also indicates that parental education is the second most important independent driver of the school track choice, while state variables like parental human capital or wealth have negligible effects, net of child skills.[43] 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 that 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, $\omega_5$. 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 that their child will receive a college taste shock that depends on their parent's education, governed by $\mu_{\nu,E^p}$. 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. However, (non-pecuniary) college costs also decrease in end-of-school skills. As derived in Section 3, for a set of children with low preschool skills, end-of-school skills are maximized if they attend the vocational school track. This force counteracts the incentive of college parents to send their child to the academic

[43] Though not directly comparable, the size of the coefficient on parental education is close to the conditional SES gap in academic school track choice found in Falk et al. (2020) or in our own data.

track described before. Third, even net of college tastes, parents have preferences regarding the school track choice directly. In particular, our calibrated parameters $\chi(E, S)$ suggest that parents bias the school track choice towards their own (college) education level.

Table 7: School Track Choice Determinants

| | Dependent Variable: $S = A$ | | | |
|---|---|---|---|---|
| | Stand. Coefficient Estimates | | | |
| | (1) | (2) | (3) | (4) |
| | Baseline | $\omega_{5,j=3,4} = 0$ | $\mu_{\nu,1} = \mu_{\nu,0} = 0$ | $\chi_0 = \chi_1 = 0$ |
| $\phi'$ | 0.09 | 0.10 | 0.10 | 0.10 |
| $\theta_3$ | 0.53 | 0.54 | 0.58 | 0.56 |
| $E = 1$ | 0.34 | 0.24 | 0.23 | 0.21 |
| $h_{11}$ | 0.00 | -0.01 | 0.00 | 0.00 |
| $a_{11}$ | 0.02 | 0.01 | 0.02 | 0.02 |

*Notes*: This table reports the standardized coefficient estimates of regressions of an academic school track dummy on all states at the time of the tracking decision. 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.

To understand how important each of these channels for the school track choice is, we perform a series of three counterfactual experiments using the calibrated model, in which we isolate each effect, respectively.[44] In particular, 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. Column (2) in Table 7 reports the (standardized) results of the regression of academic track choice on all state variables in this counterfactual scenario. The coefficient on parental education drops as expected, while the coefficient on child skills prior to the track decision increases. This confirms that the knowledge of direct parental effects on future child skill development prompts parents to send their child to the same track as their own, net of effects of parental education through child skills that are already formed.

Column (3) reports the resulting coefficient estimates when isolating the second channel, working through college tastes. If we equalize the means in college taste shocks across parental education (to zero), once again, the coefficient on direct parental influence on school

[44] 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.

track choice decreases, and the one on child skill increases. Quantitatively, these effects are comparable to the first channel. Similarly, as reported in Column (4) of Table 7, the direct influence of parental education on the school track of a child drops by almost 40% if we shut down parental preferences in school track choice directly by setting $\chi_E = 0$ for both education levels. At the same time, a child's own skills become more important for the track decision.

As discussed in Section 3, any such forces that impact 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.[45] 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. Table 8 provides an overview of selected outcomes in the baseline model (Column (1)) and compares the resulting percentage change of these outcomes in two counterfactual scenarios: In Column (2), we shut down direct parental preferences shifting the school track choice ($\chi_0 = \chi_1 = 0$) as before.[46] Moreover, in Column (3), we report the relative changes in the outcomes from another counterfactual experiment, in which 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 to be exactly the average skill level of children prior to the track choice.[47]

In both counterfactual scenarios, aggregate output in the economy increases slightly relative to the baseline economy. Note that the share of college-educated agents in the economy

[45] 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.

[46] We focus on this experiment as we view this as being the easiest to address by policies. In particular, if preferences for school tracks are coming from information frictions, as argued before, mentoring programs have proven very effective and almost cost-free in alleviating some of these frictions as argued by Falk et al. (2020) and Resnjanskij et al. (2021).

[47] 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. Picking this threshold ensures that the variance of child skills in each track is minimized. Given that the distribution of child skills is quite symmetric around its mean, this implies that approximately the bottom 50% of children in terms of their skills at the beginning of secondary school are allocated to the vocational track, while the top 50% go to the academic track.

remains constant relative to the baseline case, which is due to the general equilibrium effects on the labor market as wages adjust to keep demand for college and non-college labor approximately constant. In contrast, the share of children that attend an academic track school increases. In the case without preference-based school track choice, the share increases by 2.4%. By construction, this share increases even further in the case of the sharp track threshold, as this threshold implies that roughly 50% of the children go to either track. The reason for the positive effects on output becomes clear when we study the distribution of skills in counterfactual experiments.

In particular, the first row in Panel B. of Table 8 suggests that 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 variance in child skills within the school tracks changes relative to the baseline case, which impacts learning efficiency. In the first counterfactual experiment, skills in the academic track become more homogeneous, while the variance of skills in the vocational track increases only marginally. In the second counterfactual, the variance of skills in the vocational track decreases while the variance of skills in the academic track increases. The latter is likely due to the fact that the share of children in the academic track also increases. Overall , however, the effect on average end-of-school skills is positive, which then translates into higher output. This is consistent with the explanation of the efficiency-reducing misallocation effects that arise when parental background drives the school track choice, independently from skills.

Row 4 of Panel A. reports that without direct parental preferences in school track choice and even more so with a sharp, purely skill-based allocation rule, the dependence of school track choice on parental income decreases. Unsurprisingly, skills themselves become more important in explaining the track choice and the college choice, as shown in Row 5. However, while the intergenerational elasticity between parent's and child's income drops in the first counterfactual experiment, it slightly increases when introducing a strict skill threshold. Again, this is likely due to the fact that the share of academic track children also increases in that case.

Overall, Table 8 paints the following picture. Both counterfactual scenarios achieve an improvement in child learning during the secondary school years. This improvement yields a positive effect on aggregate output in the macroeconomy, which is larger when the track allocation is based on a pure skill threshold, though still relatively modest at 0.2%. At the same time, while decreasing parental track choice preferences improves social mobility relative to the baseline economy, this cannot be said about the case with an optimal,

Table 8: Effects of School Track Choice Counterfactuals

| Outcome | (1) Baseline Economy | (2) $\chi_0 = 0$ $\chi_1 = 0$ | (3) Sharp Skill Threshold |
|---|---|---|---|
| | | Panel A. | |
| $Y$ | 2.05 | 0.1% | 0.2% |
| College Share | 0.35 | 0.0% | 0.0% |
| A-Track Share | 0.42 | 2.4% | 10.2% |
| A-Track on Income | 0.50 | -25.2% | -38.0% |
| A-Track on Skills | 1.02 | 1.4% | 46.5% |
| CL on Skills | 0.64 | 0.4% | 3.0% |
| IGE | 0.314 | -2.2% | 1.0% |
| Gini Earnings | 0.26 | 0.4% | 0.4% |
| | | Panel B. | |
| $\bar{\theta}_5$ | 0.04 | 0.1% | 8.6% |
| $Std(\theta_{5\|S=V})$ | 0.39 | 0.1% | -2.6% |
| $Std(\theta_{5\|S=A})$ | 0.40 | -4.7% | 2.5% |

*Notes*: Column (1) shows aggregate outcomes in the baseline economy, Column (2) displays percentage changes entailed by the absence of parental preference for education, and Column (3) displays percentage changes entailed by a 50-50 rule for school tracking.

threshold-based track allocation.

### 5.3 Postponing the School Tracking Age

An important feature of school tracking policies is the age at which children are allocated across the tracks. Generally, OECD countries differ remarkably in the school tracking age (see Figure IV.2.2 in OECD (2013) for an overview). In countries with an early tracking system in place, such as Germany, it is often argued that postponing the tracking age will improve equality of opportunity in terms of access to academic education without incurring efficiency losses in terms of learning outcomes (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 of a large-scale reform that postpones the tracking age.

To evaluate such a reform in the context of Germany, we conduct a series of late tracking counterfactual experiments using our calibrated model. In each experiment, we assume that the age at which children can sort into an academic or vocational school track is postponed from 10 to 14, corresponding to model period $j = 3$ to $j = 4$. During model period $j = 3$, all children attend a school that belongs to a comprehensive school track, just like during primary school in $j = 2$. In each counterfactual experiment, all parameters, in particular those governing school track preferences and college costs, remain the same as in the baseline economy.

We present the relative changes of selected aggregate and social mobility outcomes of the counterfactual experiments relative to the baseline economy in Table 9. The experiments differ in the way we assume that prices and instruction paces are allowed to adjust. In Column (2), all prices (wages per efficiency unit for college and non-college human capital $w_0, w_1$ and the interest rate $r$) are assumed to remain at the same values as in the baseline case. That is, we compare the partial equilibrium outcomes of the policy counterfactual. Moreover, we assume that the instruction pace during the *second stage* of secondary school does not adjust. That is, the policymaker sets the same pace as in the baseline case in both academic and vocational track schools during $j = 4$. As a result, parents do not need to form expectations over the average skill levels in each track when they make the postponed track choice.

In this economy, aggregate output $Y$ is around 0.8% lower than in the baseline case. The share of college-educated agents decreases by almost 7%, and the share of children in the academic track in $j = 4$ similarly decreases by 7.4%. Average human capital is significantly less than in the baseline economy, which is ultimately a result of less efficient learning during secondary school. In particular, average end-of-school skills in period $j = 5$ are around 17% lower in the late tracking case than in the baseline economy.

As we derived in Section 3, it is theoretically not clear whether later tracking results in such learning efficiency losses. In particular, later tracking could even increase average learning outcomes if the variance of the child skill shocks is sufficiently large. The reason for that is that with large skills shocks, the gain from more homogeneous peer groups in each track during the last stage of secondary school can outweigh the losses incurred due to one more period of learning in a comprehensive track during the first stage of secondary school. However, despite sizable estimates of the child skill shocks variances, our model predicts that the learning losses from postponing four years of tracking in Germany cannot be recuperated by more-efficient learning during the remainder of secondary school.

At the same time, academic track attendance becomes less dependent on the income and education of the parent after the late tracking policy reform. For example, while in the early tracking case, 72% of children of college-educated parents go to an academic track school, this number drops to 68% in the late tracking counterfactuals. The share of children from non-college households, who attend an academic track school, however, does not drop relative to the early tracking case, resulting in effectively more children from non-academic households in academic track schools. In a similar vein, the regression coefficient of academic track attendance on parental income decreases by around 13%.

On top of that, the college decision becomes significantly less dependent on the secondary school track in the late tracking counterfactual. Concretely, while the share of academic-track graduates that go to college drops slightly, the share of vocational-track graduates going to college triples. This signals that, in the late tracking counterfactual, the benefits from academic track attendance arising from better chances to go to college are smaller than in the early tracking case. One reason for this result is that late tracking results in a less polarized distribution of end-of-school skills compared to early tracking. For example, the overall variance of end-of-school skills decreases by around 2.5%. Moreover, the difference in average skills between academic and vocational track children decreases by almost 10% in the late tracking counterfactual. Since college utility costs are decreasing in end-of-school skills, this makes the attractiveness of college education become more equal across vocational and academic school track graduates.

As a consequence, both cross-sectional inequality, as measured by the Gini coefficient of labor income, and the intergenerational elasticity of earnings decrease in the late tracking counterfactual. Thus, our quantitative exercise suggests that postponing tracking results in efficiency losses in terms of learning and aggregate output but comes with the benefit of reduced inequality and improved social mobility.

Table 9: Late Tracking Counterfactuals

| Outcome | (1) Baseline | (2) Baseline pace + PE | (3) Pace adjusts + PE | (4) Pace adjusts + GE |
|---|---|---|---|---|
| | Panel A. | | | |
| $Y$ | 2.05 | -0.8% | -0.95% | -0.2% |
| College Share | 0.35 | -6.9% | -8.1% | 0.0% |
| A-Track Share | 0.42 | -7.4% | -9.3% | -5.5% |
| CL/Non-CL Earnings | 1.773 | 0.0% | 0.0% | -0.2% |
| Gini Earnings | 0.26 | -0.4% | -0.8% | 0.0% |
| Share A-Track from CL Parent | 0.72 | -5.6% | -6.9% | -6.9% |
| Share A-Track from Non-CL Parent | 0.25 | 0.0% | 0.0% | 0.0% |
| Share CL after A-Track | 0.71 | -2.8% | -4.2% | -1.4% |
| Share CL after V-Track | 0.08 | 200% | 187% | 225% |
| A-Track on Income | 0.50 | -13.4% | -15.4% | -14% |
| IGE | 0.314 | -1.9% | -1.9% | -1.9% |
| | Panel B. | | | |
| $\bar{\theta}_5$ | 0.04 | -17.1% | -20.0% | -5.7% |
| $Std(\theta_{5\|S=V})$ | 0.39 | 1.0% | 0.8% | -0.8% |
| $Std(\theta_{5\|S=A})$ | 0.40 | -6.9% | -6.9% | -2.0% |

*Notes*: This table presents changes in outcomes due to delaying the school tracking choice by four years (from the age of ten to the age of fourteen). Column (1) shows aggregate outcomes in the baseline economy, and Columns (2) to (4) display percentage changes due to the policy change in different scenarios. Column (2): if the pace of instruction and prices are unchanged. Column (3): if the pace of instruction adjusts but prices are unchanged. Column (4): if the pace of instruction and prices adjust.

If we allow the instruction pace in each track to adjust endogenously while still keeping prices at their baseline values (in Column (3)), these two opposing effects become slightly more pronounced. For example, the share of children ending up in the academic track school drops by over 9%, and the share of college workers drops by around 8% relative to the baseline, early tracking economy. Aggregate learning also decreases more, resulting in an output loss of almost 1% in this economy. We can interpret this result again through the lens of the theoretical illustrations derived in Section 3. In particular, we have argued before that the equilibrium allocation of children across school tracks that results from a game played among parents need not be equal to the optimal one that a policymaker seeking to maximize learning would implement if there are positive direct peer effects. Against this backdrop, the results in Columns 2 and 3 of Table 9 then suggest that the unadjusted instruction pace carried over from the early tracking economy is actually closer to the one a policymaker would optimally choose than the adjusted one resulting from parents making the track decision while correctly anticipating the distribution of child skills across tracks in equilibrium. Thus aggregate learning drops. At the same time, the effects on intergenerational mobility and cross-sectional earnings inequality remain approximately the same as before.

Naturally, some of the loss in efficiency when postponing tracking could be due to the fact that, in the partial equilibrium late tracking counterfactuals, the share of college-educated workers overall declines markedly, which impacts aggregate human capital during the working years as this is assumed to grow at a college-specific rate $\gamma_{j,E}$. Once we allow for general equilibrium effects in Columns (4), the college share returns to be approximately the same as in the baseline economy, as wages adjust to clear the labor markets for college and non-college type labor. As college education becomes more attractive again, also the share of children in the academic track rises. Nonetheless, it is (at around 40%) still slightly lower than in the early tracking baseline economy. Moreover, postponing tracking still decreases the average end-of-school skills relative to early tracking (by 5.7%), yet markedly less so than in the partial equilibrium cases.

Thus, despite the fact that the variances of end-of-school skills in each school track in the late tracking counterfactual are smaller than in the baseline economy, this gain in homogeneity in peer groups cannot overcome the disadvantage in terms of average skills stemming from one more period of comprehensive track schooling. As a result, total output still decreases by 0.2% relative to the early tracking economy.

On the other hand, the resulting general equilibrium continues to feature more mobility between generations, as the intergenerational earnings elasticity drops by almost 2%. This is again a consequence of significantly more children going to college after a vocational track secondary school and a declining share of children from academic parents going to academic track schools. Perhaps surprisingly, cross-sectional inequality as measured by the Gini coefficient on labor income, does not change relative to the early tracking case.

The main takeaways of the policy reform that postpones school tracking to age 14 in our model can be summarized as follows. First, postponing school tracking incurs efficiency losses from worse learning outcomes in the additional period of comprehensive school. The losses cannot be compensated by gains in later years that arise from more homogeneous peer groups across tracks as the track decision is based on more complete information about children's skill evolution. Second, later tracking incentivizes fewer parents to send their child to an academic track secondary school as the likelihood of college education depends less on the secondary school track. Third, this results in more equal access to academic secondary education by parental background, which leads to more equal access to higher education and more equal labor market outcomes. The quantitative size of this possible efficiency-equity trade-off depends on whether the school tracking age reform is evaluated in the short run, when wages and possibly the instruction paces in schools have not reacted, or in the long run, when general equilibrium effects are taken into account. Finally, in all cases, later tracking reduces the persistence of economic status across generations.

# 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 many advanced countries, influences not only equality of educational opportunities of children from different parental backgrounds, but also shapes aggregate learning and, as a consequence, aggregate economic efficiency. We add a macroeconomic perspective to the predominantly reduced-form literature by building a macroeconomic GE model of overlapping generations that specifically zooms in on the schooling years of the children. To that end, we formulate a simple theory of child skill formation, where child skills depend linearly on her classroom peers and non-linearly on the instruction pace that is specific to each school track.

We show that this child skill formation technology alone entails theoretical implications for the effect of school tracking policies on the distribution of child skills that are in line with the most robust findings of a vast empirical literature as well as the most popular arguments in the public debate about tracking. In particular, not every child gains from tracking, and the losses are often concentrated among lower-skilled children. Additionally, tracking can lead to increased inequality in end-of-school skills. Finally, the effects of tracking on learning efficiency, while typically positive on average, depend on the age at which children are tracked and the size of uncertainty regarding the evolution of child skills, highlighting the importance of the timing of tracking.

We embed this theory into a standard Aiyagari-style life-cycle framework in which parents make a school track decision for their children. We tailor the model to fit the German Education System, where the track decision occurs at the age of 10 of the child, and calibrate it on German data. Our quantitative results suggest that the skills accumulated during secondary schooling are a major contributor to lifetime inequality and that variation coming from the initial school track alone can account for around 12% of the variation in eventual lifetime earnings. Conditional on prior child skills, the track choice is strongly influenced by parental preferences that cannot be explained by parental inputs into child skills or tastes for higher education. This gives rise to efficiency-reducing misallocation of children across tracks. Our results indicate that policies that reduce the parental influence on the school track choice, such as mentoring policies (Falk et al., 2020), can, therefore, not only improve social mobility but also lead to modest efficiency gains in terms of aggregate output in the macroeconomy.

Our paper also shows 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 0.2% of GDP while decreasing the intergenerational income elasticity by around 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 from the track decision depending less on the parental background and the college decision depending less on the secondary school track.

# References


Aakvik, A., Salvanes, K. G., and Vaage, K. (2010). Measuring heterogeneity in the returns to education using an education reform. *European Economic Review*, 54(4):483–500.

Abbott, B., Gallipoli, G., Meghir, C., and Violante, G. L. (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy*, 127(6):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., Doepke, M., Sorrenti, G., and Zilibotti, F. (2023). It takes a village: the economics of parenting with neighborhood and peer effects. Working Paper w27050, National Bureau of Economic Research.

Agostinelli, F., Saharkhiz, M., and Wiswall, M. (2019). Home and school in the development of children. Working Paper w26037, National Bureau of Economic Research.

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

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

Becker, G. S. and Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of labor economics*, 4(3, Part 2):S1–S39.

Bellenberg, G. and Forell, M. (2012). Schulformwechsel in deutschland. *Durchlässigkeit und Selektion in den*, 16.

Blossfeld, H., Roßbach, H., and von Maurice, J. (2011). The german national educational panel study (neps). *Zeitschrift für Erziehungswissenschaft: Sonderheft*, 14.

Bonesrønning, H., Finseraas, H., Hardoy, I., Iversen, J. M. V., Nyhus, O. H., Opheim, V., Salvanes, K. V., Sandsør, A. M. J., and Schøne, P. (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., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of human Resources*, 43(4):972–1059.

Capelle, D. (2022). The great gatsby goes to college: Tuition, inequality and intergenerational mobility in the u.s. Working paper.

Caucutt, E. M. and Lochner, L. (2020). Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy*, 128(3):1065–1147.

Ciccone, A. and Peri, G. (2005). Long-run substitutability between more and less educated workers: evidence from us states, 1950–1990. *Review of Economics and statistics*, 87(4):652–663.

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

Cunha, F., Heckman, J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Daruich, D. (2022). The macroeconomic consequences of early childhood development policies. Working Paper 2018-29, FRB St. Louis.

Dodin, M., Findeisen, S., Henkel, L., Sachs, D., and Schüle, P. (2021). Social mobility in germany. Discussion Paper DP16355, CEPR.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, 101(5):1739–74.

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

Dustmann, C., Puhani, P. A., and Schönberg, U. (2017). The long-term effects of early track choice. *The Economic Journal*, 127(603):1348–1380.

Epple, D. and Romano, R. (2011). Peer effects in education: A survey of the theory and evidence. In *Handbook of social economics*, volume 1, pages 1053–1163. Elsevier.

Falk, A., Kosse, F., and Pinger, P. (2020). Mentoring and schooling decisions: Causal evidence. *Journal of Political Economy (forthcoming).*

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

Fuchs-Schündeln, N., Krueger, D., Ludwig, A., and Popova, I. (2022). The long-term distributional and welfare effects of covid-19 school closures. *The Economic Journal*, 132(645):1647–1683.

Fujimoto, J., Lagakos, D., and Vanvuren, M. (2023). Aggregate and distributional effects of 'free' secondary schooling in the developing world. Working Paper w31029, National Bureau of Economic Research.

Guyon, N., Maurin, E., and McNally, S. (2012). The effect of tracking students by ability into different schools a natural experiment. *Journal of Human resources*, 47(3):684–721.

Hanushek, E. A. and Wössmann, L. (2006). Does educational tracking affect performance and inequality? differences-in-differences evidence across countries. *The Economic Journal*, 116(510):C63–C76.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Heckman, J. and Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, 6(1):689–733.

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

Huggett, M., Ventura, G., and Yaron, A. (2011). Sources of lifetime inequality. *American Economic Review*, 101(7):2923–54.

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

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of political Economy*, 105(3):473–522.

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

Kyzyma, I. and Groh-Samberg, O. (2018). "intergenerational economic mobility in germany: Levels und trends ". Working paper, DIW.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849.

Lee, S. Y. and Seshadri, A. (2019). On the intergenerational transmission of economic status. *Journal of Political Economy*, 127(2):855–921.

Malamud, O. and Pop-Eleches, C. (2011). School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics*, 95(11-12):1538–1549.

Matthewes, S. H. (2021). Better together? heterogeneous effects of tracking on student achievement. *The Economic Journal*, 131(635):1269–1307.

Meghir, C. and Palme, M. (2005). Educational reform, ability, and family background. *American Economic Review*, 95(1):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*.

Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E., and Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for educational research online*, 5(2):80–109.

OECD (2013). Pisa 2012 results: What makes schools successful (volume iv): Resources, policies and practices, pisa. Report TD/TNC 114.1481, OECD.

Passaretta, G., Skopek, J., and van Huizen, T. (2022). Is social inequality in school-age achievement generated before or during schooling? a european perspective. *European Sociological Review*, 38(6):849–865.

Pekkala Kerr, S., Pekkarinen, T., and Uusitalo, R. (2013). School tracking and development of cognitive skills. *Journal of Labor Economics*, 31(3):577–602.

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.

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

Pohl, S. and Carstensen, C. H. (2013). Scaling of competence tests in the national educational panel study-many questions, some answers, and further challenges. *Journal for Educational Research Online*, 5(2):189–216.

Resnjanskij, S., Ruhose, J., Wiederhold, S., and Woessmann, L. (2021). Can mentoring alleviate family disadvantage in adolscence? a field experiment to improve labor-market prospects.

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

Schneider, T. and Linberg, T. (2022). Development of socio-economic gaps in children's language skills in germany. *Longitudinal and Life Course Studies*, 13(1):87–120.

Woessmann, L. (2013). Die entscheidende säule. *Wirtschaftswoche Global*, 2(24.06.2013):110–111.

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

Yum, M. (2022). Parental time investment and intergenerational mobility. *International Economic Review*.

# A Proof of Propositions

## Proposition 1

For the proof of this proposition, we denote by $\theta_1$ the child skills at the beginning of secondary school and by $\theta_2$ the skills at the end of secondary school. First, we show that maximizing the aggregate end-of-school skills in a tracking system implies a threshold skill level $\tilde{\theta}_1$, such that all $\theta_1 < \tilde{\theta}_1$ go to one track, call it $S = V$ and all $\theta_1 > \tilde{\theta}_1$ go to the other track, $S = A$ (and those with $\theta_1 = \tilde{\theta}_1$ 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 $\theta_2$ in (3) of a child in a given school track $S$ with instruction pace $P_S^*$ using Lemma 1 as:

$$\theta_2 = \theta + \alpha\bar{\theta}_S + \frac{\beta^2}{2\delta} + \frac{\beta\gamma\theta_1}{\delta} + \frac{\gamma^2\theta_1\bar{\theta}_S}{\delta} - \frac{\gamma^2\bar{\theta}_S^2}{2\delta} + \eta_2. \tag{A.1}$$

After adding and subtracting $\frac{\gamma^2}{2\delta}\theta_1^2$, this can be expressed as

$$\begin{aligned}\theta_2 &= \theta_1 + \alpha\bar{\theta}_S + \frac{\beta^2}{2\delta} + \frac{\beta\gamma\theta_1}{\delta} + \frac{\gamma^2\theta_1^2}{2\delta} + \eta_2 - \frac{\gamma^2}{2\delta}\left(\theta_1^2 - 2\theta_1\bar{\theta}_S + \bar{\theta}_S^2\right)\\ &= \theta_2(P_{\theta_1}^*) - \frac{\gamma^2}{2\delta}(\theta_1 - \bar{\theta}_S)^2,\end{aligned} \tag{A.2}$$

where $\theta_2(P_{\theta_1}^*)$ denotes end-of-school skills if the child with skills $\theta_1$ is taught at her individually optimal teaching pace $P_{\theta_1}^*$. Thus, in a given track, end-of-school skills are a strictly decreasing function of the *distance* to the average skill level $\bar{\theta}_S$ 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 $\mathbb{E}[\theta_2]$ 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, $\mathbb{E}(\theta_1|S = V) < \mathbb{E}(\theta_1|S = A)$. Then, because there is no strict threshold, this means that for any initial skill level $\theta_1$, there must be at least two children with initial skill levels smaller or equal to $\theta_1$ that go to different tracks or at least two children with initial skill levels larger or equal than $\theta_1$ that go to different tracks. This implies that there exists a child with $\theta_1' \leq \mathbb{E}(\theta_1|S = V)$ that goes to track $S = A$, and/or a child with

$\theta_1' \geq \mathbb{E}(\theta_1|S = A)$ that goes to track $S = V$, and/or two children with skills $\theta_1' < \theta_1''$, with $\theta_1', \theta_1'' \in [\mathbb{E}(\theta_1|S = V), \mathbb{E}(\theta_1|S = A)]$, 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 $\theta_1'$ 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 $\tilde{\theta}_1$ be the skill threshold and let $S$ again indicate to which track a child is allocated, now with $S = V$ for all $\theta_1 \leq \tilde{\theta}_1$ and $S = A$ for all $\theta_1 > \tilde{\theta}_1$.

A policymaker solves

$$\begin{aligned} & \max_{\tilde{\theta}_1} \quad \mathbb{E}(\theta_2) \\ \iff & \max_{\tilde{\theta}_1} \quad \mathbb{E}(\mathbb{E}(\theta_2|S)) \\ & \text{subject to} \\ & \qquad P_S \text{ chosen optimally given Lemma 1.} \end{aligned} \tag{A.3}$$

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

$$\begin{aligned} & \max_{\tilde{\theta}_1} \quad \frac{\beta^2}{2\delta} + \frac{\gamma^2}{2\delta}\, \mathbb{E}\left(\mathbb{E}(\theta_1|S)^2\right) \\ \iff & \max_{\tilde{\theta}_1} \quad \frac{\beta^2}{2\delta} + \frac{\gamma^2}{2\delta} \left( F(\tilde{\theta}_1)\, \mathbb{E}(\theta_1|\theta_1 \leq \tilde{\theta}_1)^2 + (1 - F(\tilde{\theta}_1))\, \mathbb{E}(\theta_1|\theta_1 > \tilde{\theta}_1)^2 \right), \end{aligned} \tag{A.4}$$

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 $\tilde{\theta}_1$. Using the law of total variance, the maximization problem can thus be rewritten as (dropping the constant term)

$$\begin{aligned} &\max_{\tilde{\theta}_1} \quad \mathbb{E}(\theta_2) \\ \iff &\max_{\tilde{\theta}_1} \quad \frac{\gamma^2}{2\delta}\left(\sigma^2_{\theta_1} - \mathbb{E}(Var[\theta_1|S])\right). \end{aligned} \tag{A.5}$$

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 $\tilde{\theta}_1^* = \mathbb{E}\,\theta_1 = 0$, 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 $\hat{\theta}_1$ is indifferent between tracks $V$ and $A$ iff

$$\begin{aligned} &\left(\alpha + \hat{\theta}_1\frac{\gamma^2}{\delta}\right)\mathbb{E}(\theta_1|\theta_1 \le \hat{\theta}_1) - \frac{\gamma^2}{2\delta}\,\mathbb{E}(\theta_1|\theta_1 \le \hat{\theta}_1)^2 \\ = &\left(\alpha + \hat{\theta}_1\frac{\gamma^2}{\delta}\right)\mathbb{E}(\theta_1|\theta_1 > \hat{\theta}_1) - \frac{\gamma^2}{2\delta}\,\mathbb{E}(\theta_1|\theta_1 > \hat{\theta}_1)^2 \\ \iff &\left(-\alpha - \hat{\theta}_1\frac{\gamma^2}{\delta}\right)\sigma_{\theta_1}\frac{f(\hat{\theta}_1/\sigma)}{F(\hat{\theta}_1/\sigma)} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_1}\frac{f(\hat{\theta}_1/\sigma)^2}{F(\hat{\theta}_1/\sigma)^2} \\ = &\left(\alpha + \hat{\theta}_1\frac{\gamma^2}{\delta}\right)\sigma_{\theta_1}\frac{f(\hat{\theta}_1/\sigma)}{1 - F(\hat{\theta}_1/\sigma)} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_1}\frac{f(\hat{\theta}_1/\sigma)^2}{(1 - F(\hat{\theta}_1/\sigma))^2} \end{aligned} \tag{A.6}$$

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 $\hat{\theta}_1$ 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 $\hat{\theta}_1 = 0$, as can be directly seen from (A.6). When $\alpha > 0$, the root is smaller than 0, i.e. $\hat{\theta}_1 < 0$.

## 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 $\theta_2$ is monotonically increasing in $\theta_1$ in a given track (see (A.1)), this is the child with the highest initial skill, say $\theta_{1,max}$. Moreover, from the properties of a truncated normal distribution, we know that, for any skill threshold $\tilde{\theta}_1$, average skills in the $A$ track, $\bar{\theta}_{1,A}$ are larger than the unconditional average, $\bar{\theta}_{1,C} = 0$. Thus, the squared distance between $\theta_{1,max}$ and $\bar{\theta}_{1,A}$ in a tracking system is smaller. Taken together, (A.2) implies that the child with initial skill $\theta_{1,max}$ 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 $\theta_2$ are monotonically increasing in $\theta_1$ in every track, the range is characterized by the intersection of the linear function $\theta_{2,C}(\theta_1, \bar{\theta}_{1,C})$ with $\theta_{2,V}(\theta_1, \bar{\theta}_{1,V})$ and $\theta_{2,A}(\theta_1, \bar{\theta}_{1,A})$. For any skill threshold, the lower intersection $\theta_{1,L}$ hence solves

$$\begin{aligned} & \theta_{1,L} + \alpha\bar{\theta}_{1,C} + \frac{\beta^2}{2\delta} + \frac{\beta\gamma}{\delta}\theta_{1,L} + \frac{\gamma^2}{\delta}\bar{\theta}_{1,C}\theta_{1,L} - \frac{\gamma^2}{2\delta}\bar{\theta}^2_{1,C} + \eta_2 \\ = \quad & \theta_{1,L} + \alpha\bar{\theta}_{1,V} + \frac{\beta^2}{2\delta} + \frac{\beta\gamma}{\delta}\theta_{1,V} + \frac{\gamma^2}{\delta}\bar{\theta}_{1,V}\theta_{1,L} - \frac{\gamma^2}{2\delta}\bar{\theta}^2_{1,V} + \eta_2 \\ \iff \quad \theta_{1,L} = \quad & \frac{1}{2}\bar{\theta}_{1,V} - \frac{\alpha\delta}{\gamma^2}. \end{aligned} \tag{A.7}$$

Similarly, the upper intersection is given at

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

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

If $\alpha > 0$, and the policymaker enforces the tracking skill threshold $\tilde{\theta}_1 = 0$, 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 $[\theta_{1,L}, 0]$ is allocated into the $V$ track but loses relative to a comprehensive system. Similarly, every child with an initial skill inside $[0, \theta_{1,U}]$ is allocated to track $A$ but loses relative to a comprehensive system. With $\alpha > 0$, $|\theta_{1,U}| < |\theta_{1,L}|$ and hence, the range of children in the $A$ track that lose is smaller. The interval $[0, \theta_{1,U}]$ 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 $\theta_1$ the child skills at the beginning of secondary school, by $\theta_2$ the skills at the intermediary stage of secondary school and by $\theta_3$ the skills at the end of secondary school. First, we characterize the variance of $\theta_2$. We start by collecting expressions for conditional and unconditional first and second moments.

The unconditional expected value of $\theta_2$ in track $V$, if everyone went to $V$ is

$$\begin{aligned} \mathbb{E}(\theta_{2,V}) &= \frac{\beta^2}{2\delta} + \alpha\bar{\theta}_{1,V} - \frac{\gamma^2}{2\delta}\bar{\theta}^2_{1,V} \\ &= \frac{\beta^2}{2\delta} - \alpha\sigma_{\theta_1}\frac{f(\tilde{\theta}_1/\sigma_{\theta_1})}{F(\tilde{\theta}_1/\sigma_{\theta_1})} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_1}\frac{f(\tilde{\theta}_1/\sigma_{\theta_1})^2}{F(\tilde{\theta}_1/\sigma_{\theta_1})^2}. \end{aligned} \tag{A.9}$$

The unconditional expected value of $\theta_2$ in track $A$, if everyone went to $A$ is

$$\begin{aligned} \mathbb{E}(\theta_{2,A}) &= \frac{\beta^2}{2\delta} + \alpha\bar{\theta}_{1,A} - \frac{\gamma^2}{2\delta}\bar{\theta}^2_{1,A} \\ &= \frac{\beta^2}{2\delta} + \alpha\sigma_{\theta_1}\frac{f(\tilde{\theta}_1/\sigma_{\theta_1})}{1 - F(\tilde{\theta}_1/\sigma_{\theta_1})} - \frac{\gamma^2}{2\delta}\sigma^2_{\theta_1}\frac{f(\tilde{\theta}_1/\sigma_{\theta_1})^2}{(1 - F(\tilde{\theta}_1/\sigma_{\theta_1}))^2}. \end{aligned} \tag{A.10}$$

The variance of $\theta_2$ in a comprehensive system is

$$\begin{aligned} Var(\theta_{2,C}) &= \mathbb{E}\big((\theta_2 - \mathbb{E}(\theta_2))^2\big) \\ &= (1+\beta)^2\sigma^2_{\theta_1} + \sigma^2_{\eta 2} \\ &\sigma^2_{\theta_2,C} + \sigma^2_{\eta 2}, \end{aligned} \tag{A.11}$$

where we define $\sigma^2_{\theta_2,C}$ to be the variance of $\theta_2$ 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

$$
\begin{aligned}
\mathbb{E}(\theta_{3,LT}) &= \mathbb{E}(\mathbb{E}(\theta_{3,LT}|S_{LT}^2)) \\
&= \mathbb{E}(\theta_{2,LT}) + \frac{\beta^2}{2\gamma} + (\alpha+\beta)\,\mathbb{E}(\mathbb{E}(\theta_{2,LT}|S_{LT})) + \frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_{2,LT}|S_{LT})^2) \\
&= (2+\alpha+\beta)\frac{\beta^2}{2\gamma} + \frac{\gamma}{2}[\sigma^{\theta_2,LT} - \mathbb{E}(Var(\theta_{2,LT}|S_{LT}))],
\end{aligned} \tag{A.12}
$$

where $\mathbb{E}(\theta_{2,LT})$ and $\sigma^2_{\theta_2,LT}$ 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 $\theta_{2,LT} \leq \tilde{\theta}_{2,LT}$ and $S_{LT} = A$ otherwise. The cut-off that maximizes (A.12) is $\tilde{\theta}^*_{2,LT} = \mathbb{E}(\theta_{2,LT}) = \frac{\beta^2}{2\gamma}$. 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 $\theta_{3,LT}$ 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

$$
\begin{aligned}
\mathbb{E}(\theta_{3,ET}) &= \mathbb{E}(\mathbb{E}(\theta_{3,ET}|S_{ET}^2)) \\
&= \frac{\beta^2}{2\gamma} + (1+\alpha+\beta)\,\mathbb{E}(\mathbb{E}(\theta_{2,ET}|S_{ET})) + \beta\frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_{2,ET}|S_{ET})^2) \\
&= \frac{\beta^2}{2\gamma} + (1+\alpha+\beta)\left(\frac{\beta^2}{2\gamma} + \beta\frac{\gamma}{2}[\sigma^2_{\theta_1} - \mathbb{E}(Var(\theta_{1,ET}|S_{ET}))]\right) + \beta\frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_{2,ET}|S_{ET})^2 \\
&= \frac{\beta^2}{2\gamma} + (1+\alpha+\beta)\left(\frac{\beta^2}{2\gamma} + \beta\frac{\gamma}{2}[\sigma^2_{\theta_1} - \mathbb{E}(Var(\theta_{1,ET}|S_{ET}))]\right) \\
&+ \beta\frac{\gamma}{2}[\sigma^2_{\theta_2,ET} - \mathbb{E}(Var(\theta_{2,ET}|S_{ET}))].
\end{aligned} \tag{A.13}
$$

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

$$
\begin{aligned}
&\mathbb{E}(\theta_{3,LT}) - \mathbb{E}(\theta_{3,ET}) \\
&= \beta\frac{\gamma}{2}\left(\mathbb{E}(\mathbb{E}(\theta_{2,LT}|S_{LT})^2) - \mathbb{E}(\mathbb{E}(\theta_{2,ET}|S_{ET})^2)\right) \\
&- (1+\alpha+\beta)\beta\frac{\gamma}{2}\,\mathbb{E}(\mathbb{E}(\theta_1|S_{ET})^2) > 0.
\end{aligned} \tag{A.14}
$$

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 $\theta_2$ in a late tracking system is given by

$$\mathbb{E}(\theta_{2,LT}|S_{LT}=V)=\frac{\beta^2}{2\gamma}-\sigma_{\theta_2,LT}\frac{f(\tilde{\theta}_{2,LT}/\sigma_{\theta_2,LT})}{F(\tilde{\theta}_{2,LT}/\sigma_{\theta_2,LT})} \tag{A.15}$$

and

$$\mathbb{E}(\theta_{2,LT}|S_{LT}=A)=\frac{\beta^2}{2\gamma}+\sigma_{\theta_2,LT}\frac{f(\tilde{\theta}_{2,LT}/\sigma_{\theta_2,LT})}{1-F(\tilde{\theta}_{2,LT}/\sigma_{\theta_2,LT})}, \tag{A.16}$$

where the unconditional variance of $\theta_2$ in a late tracking system is given by $\sigma^2_{\theta_2,LT}=\sigma^2_{\theta_2,C}+\sigma^2_{\eta_2}$, i.e. by the one computed in equation (A.11). Since late tracking occurs *after* the realization of skill shocks in period 2, 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 $\tilde{\theta}_1=\mathbb{E}(\theta_1)=0$ and $\tilde{\theta}_2=\mathbb{E}(\theta_{2,LT})=\frac{\beta^2}{2\gamma}$. In that case, we can write the expressions for the various expected square conditional expected values as follows:

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(\theta_1|S_{ET})^2)&=2\chi\sigma^2_{\theta_1}\\
\mathbb{E}(\mathbb{E}(\theta_{2,LT}|S_{LT})^2)&=\frac{\beta^4}{4\gamma^2}+2\chi(\sigma^2_{\theta_{2,LT}}+\sigma^2_{\eta_2})\\
\mathbb{E}(\mathbb{E}(\theta_{2,ET}|S_{ET})^2)&=\frac{\beta^4}{4\gamma^2}+2\chi\sigma^2_{\theta_1}\left(\alpha^2+\gamma^2f(0)^2\sigma^2_{\theta_1}-\frac{\beta^2}{2}\right)\\
+2f(0)\sigma^2_{\theta_1}&\left(\beta^2+2\alpha(1+\beta)-(2\gamma f(0)\sigma_{\theta_1})^2\right)+2\chi(\sigma^2_{\theta_{2,LT}}+2\chi\gamma^2\sigma^2_{\theta_1}).
\end{aligned}$$

Condition (A.14) then becomes

$$\begin{aligned}
&\mathbb{E}(\theta_{3,LT})-\mathbb{E}(\theta_{3,ET})\\
&=\beta\frac{\gamma}{2}\Bigg(2\chi\sigma^2_{\eta_2}-2\chi\sigma^2_{\theta_1}\Bigg(\alpha^2+\gamma^2f(0)^2\sigma^2_{\theta_1}-\frac{\beta^2}{2}\\
&+\beta^2+2\alpha(1+\beta)-4\gamma^2f(0)^2\sigma^2_{\theta_1}+2\chi\gamma^2\sigma^2_{\theta_1}+1+\alpha+\beta\Bigg)\Bigg)\\
&=\frac{\gamma}{\pi}\left(\sigma^2_{\eta_2}-\sigma^2_{\theta_1}\left(1+\alpha+\alpha^2+\beta+\frac{\beta^2}{2}+2\alpha(1+\beta)+\frac{\gamma^2}{2\pi}\sigma^2_{\theta_1}\right)\right)>0.
\end{aligned} \tag{A.17}$$

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. it equates marginal products to prices.

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 market for each education level clears.
   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.
   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

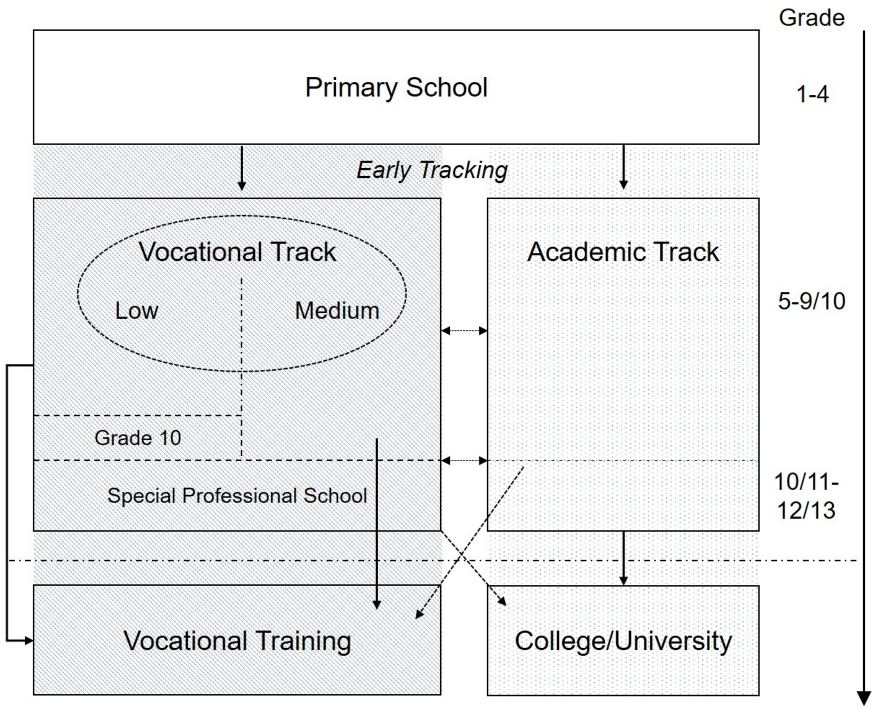
In this section, we provide an overview of the most important features of the German Education and School System. A more extensive description can be found, for example, in Henninges et al. (2019). Figure C.1 illustrates a simplified structure of the system, starting in Grade 4 and ending with tertiary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

Generally, schooling is mandatory in Germany for every child starting at age six and lasting for nine or ten years. At age six, all children visit a comprehensive primary school that lasts the first four grades.[48] 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.

[48] 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




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 for academic track schools. However, regardless of the secondary school track, becoming a teacher requires 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.

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)

# 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 (Neumann et al., 2013). To that end, the NEPS carefully designs and implements regular tests of the respondents' competencies along several domains. Given its central role not only in educational contexts but also as a predictor for later labor market success, we focus on mathematical competencies. Following the guidelines set by the Program for International Student Assessment (PISA), the mathematical competence domain is not just 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 mathematics to constructively engage with real-life problems (Neumann et al., 2013). The test, therefore, 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 primarily simple and complex multiple-choice questions, as well as short-constructed responses.[49]

In order to use these questions for the analysis of latent competencies, they need to be scaled. The NEPS (similar to the PISA) uses a scaling procedure that follows 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

[49] 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). The mathematical competence test primarily consists of simple multiple-choice questions.

the 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 most important scaling model used by the NEPS is the Rasch model. 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 $\theta_v$ denotes the latent ability of person $v$ and $\sigma_i$ 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.[50] This model is estimated via (weighted) conditional maximum likelihood under a normality assumption on latent ability.

Table D.1 describes NEPS samples of mathematics assessments by cohort and Grade level.

# E 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. For that we use data from the NEPS Starting Cohort 3. For this cohort, we have information on the school track a child attended in grade 5 as well as the school track that was recommended to the child by her primary school teacher. Moreover, we have information on the highest education level of the parents. We define the dummy variable High SES as being one whenever at least one of the parents obtained college education. Finally, we have information about the test scores each child achieved at the very beginning of secondary school (in grade 5) and towards the end of secondary school (grade 9). See also Section D for more details on the test scores.

The left panel of Table E.1 shows the results of a regression an academic track dummy on the High SES dummy, controlling for age and gender of the child. In the first column, we do not control for math test scores at the beginning of secondary school. The estimated coefficient of 0.35 suggests that, unconditionally, children from parents with a college back-

[50] It is interval-scaled asBallou (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

ground are 35 percentage points more likely to go to an academic track school. In the second column, we control for the math test scores of the child. Even then, children from a higher education family background are significantly more likely to go to an academic track school. The conditional SES gap in academic track attendance is 24 percentage points. Both the unconditional and the conditional SES gap in the NEPS data are consistent with other estimates from Germany, such as Falk et al. (2020).

The right panel compares the reaction of high and low SES parents to the track recommendations of their child. For example, the top two rows suggest that, after having received an academic track recommendation, almost 95% of high SES parents follow that recommendation and send their child to an academic track. However, among low SES parents, this share is only 81%. Thus, a significant portion of low SES parents deviate downwards from the recommended track. Vice versa, after receiving a vocational track recommendation, more than 20% of high SES parents deviate from that recommendation and send their child the academic track regardless. This share is less than 10% for the low SES parents.

In Table E.2, we present the results of a linear regression of test scores (math and reading) in grade 9 on test scores at the beginning of secondary school (grade 5), an academic track dummy, a parental high SES dummy and dummies that equal one if the track choice deviated up (that is from vocational recommendation to academic actual track) or down (from academic recommendation to

Table E.1: Parental Influence in School Track Selection

| | Academic Track | |
|---|---|---|
| High SES | 0.35 | 0.24 |
| | (0.02) | (0.02) |
| Controls: | | |
| Age & Gender | yes | yes |
| Tests | no | yes |
| $R^2$ | 0.2 | 0.36 |
| $N$ | 2,480 | 2,475 |

| Deviations from Teacher Recom. | | |
|---|---|---|
| | High SES | Low SES |
| **Academic Recom.** | | |
| Follow | 94% | 81% |
| Deviate | 6% | 19% |
| **Vocational Recom.** | | |
| Follow | 78% | 91% |
| Deviate | 22% | 9% |

vocational actual track). The regressions also control for age and gender of the child.

The estimated coefficient on the upward deviation are negative and statistically significant. The size of the coefficient suggest that, when math score is the dependent variable, having deviated up undoes around half of the learning game that comes from being in an academic track. In terms of reading outcomes, the deviation up dummy even completely undoes the gains associated from being in an academic track.

Similarly, the coefficient on downward deviation is positive. It is statistically significant only when reading scores are the dependent variable. Its size suggests that less than half of the average loss from not being in an academic track school is undone because the child deviated downward.

In sum, these results suggest firstly that track recommendations are, on average, reasonable. That is because deviations from the recommendations are harmful in terms of learning outcomes of the child in case of downward deviations and inconsequential for children in case of upward deviations.

# F Details on Child Skill Technology Estimation

We employ a linear measurement system for the logarithm of latent skills in each period that is given by

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

where $M_{i,k,j}$ denotes the $k$th 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 and are discussed in detail below. 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

Table E.2: Effect of Deviations on Learning

| | Dependent Variable | |
|---|---|---|
| Coefficient | Math Score (Grade 9) | Reading Score (Grade 9) |
| Score (Grade 5) | 0.574 | 0.427 |
| | (0.023) | (0.020) |
| Ac. Track | 0.101 | 0.106 |
| | (0.011) | (0.010) |
| Deviation Up | -0.054 | -0.100 |
| | (0.019) | (0.020) |
| Deviation Down | 0.027 | 0.040 |
| | (0.018) | (0.020) |
| High SES | 0.028 | 0.028 |
| | (0.009) | (0.009) |
| $R^2$ | 0.456 | 0.387 |
| $N$ | 1,904 | 1,816 |

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

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

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

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

Equipped with identified latent variables up to measurement error for all periods, we can plug these into the child skill technology (28), which yields

[51] 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.

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

where $\overline{\overline{\widetilde{M}}}_{-i,j,S}$ refers to the expected value of the $k$th transformed measure across all children other than $i$ in a classroom in track $S$ and $\overline{\overline{\widetilde{M}}}_{j,S}$ to that of the expected 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 $\eta$ conditional on all independent variables holds, an OLS estimator of (F.4) will be biased. To account for that, we follow the literature and Bartlett factors scores to to aggregate the different measures into an unbiased score

Tables F.1 and F.2 describe the evolution of child skills over time using the identified latent variables.

Table F.1: Differences in Average Skills in Standard Deviation

| | | Difference by | |
|---|---|---|---|
| | Grade | Parent's Education | School Track |
| Cohort 2 | G1 | 0.46 | 0.77 |
| | G4 | 0.55 | 0.83 |
| | G7 | 0.45 | 0.93 |
| Cohort 3 | G5 | 0.48 | 0.87 |
| | G7 | 0.55 | 0.92 |
| | G9 | 0.57 | 1.01 |
| Cohort 4 | G9 | 0.54 | 0.97 |

*Notes*: This table provides information on average differences identified latent math grades up to measurement error in one standard deviation unit by parental background and school track over time. All observations are weighted. Source: NEPS

Table F.2: Rank-Rank Correlations

| | | Rank-Rank Correlation | Obs. |
|---|---|---|---|
| Cohort 2 | G1-G4 | 0.57 | 3,116 |
| | G4-G7 | 0.61 | 2,267 |
| Cohort 3 | G5-G9 | 0.67 | 4,927 |

*Notes*: This table provides the rank-rank correlations in identified latent math grades up to measurement error. Source: NEPS