European University Institute
DEPARTMENT OF POLITICAL AND SOCIAL SCIENCES

Cracking Meritocracy from the Starting Gate

Social Inequality in Skill Formation and School Choice

Carlos J. Gil Hernández

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Political and Social Sciences
of the European University Institute

Florence, 29 October 2020

European University Institute
Department of Political and Social Sciences

Cracking Meritocracy from the Starting Gate

Social Inequality in Skill Formation and School Choice

Carlos J. Gil Hernández

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Political and Social Sciences
of the European University Institute

Examining Board

Professor Fabrizio Bernardi, European University Institute (EUI Supervisor)
Professor Juho Härkönen, European University Institute
Professor Jonas Radl, Carlos III University / WZB Berlin Social Science Center
Professor Leire Salazar, Universidad Nacional de Educación a Distancia (UNED)

Researcher declaration to accompany the submission of written work Department of Political and Social Sciences - Doctoral Programme

I Carlos J. Gil Hernández certify that I am the author of the work "Cracking Meritocracy from the Starting Gate. Social Inequality in Skill Formation and School Choice" I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297).

I declare that this work consists of 85,083 words.

Statement of inclusion of previous work:

I confirm that Chapter IV was partially the result of a previous research fellowship funded by the European Consortium for Sociological Research Visitor Grant I undertook at the WZB Berlin Social Science Center.

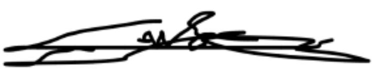
I confirm that Chapter II was jointly co-authored with Mr Marco Cozzani and Mr Fabrizio Bernardi and I contributed 60% of the work.

I confirm that Chapter III draws upon an earlier article I published in the journal *Sociology of Education*: *https://doi.org/10.1177/0038040719830698*

Statement of language correction:

This thesis has been corrected for linguistic and stylistic errors. I certify that I have checked and approved all language corrections, and that these have not affected the content of this work.

Signature and date:



09/10/2020

Abstract

In post-industrial societies, a college education is the main channel for upper classes to prevent their children falling down the social ladder, while, for working classes, it is the best bet for upward mobility. Despite attaining post-compulsory education was equalised and a driver of social mobility in the last decades, inequalities by socioeconomic status (SES) in college graduation, the main social lift, remained relatively unchanged. We are only starting to understand the complex interplay between biological and environmental factors explaining why educational inequalities gestate before birth and persist over generations. Besides, further research is needed to unravel why advantaged students are more likely to get ahead in education than equally-skilled, but disadvantaged peers.

This thesis bridges interdisciplinary literature to study how parental SES affects educational attainment during childhood in Germany, evaluating the implications for social justice. It contributes to the literature by (1) analysing the consequences of prenatal health shocks on skill formation; (2) examining the effect of cognitive and non-cognitive skills on the transition to secondary education; and (3) assessing SES-heterogeneity in these associations. Drawing from *compensatory theories*, I demonstrate how negative traits for educational attainment—low birth weight and cognitive ability—are less detrimental for high-SES children from the early stages of the status-attainment process due to mechanisms like parental investments and aspirations, and teachers' bias in assessments.

The German educational system enforces early tracking into academic or vocational pathways from age 10, supposedly according to ability. Thus, the case of Germany represents an institutional starting gate to evaluate equal opportunity, where compensating for negative traits might be difficult. To test *compensatory theories*, I utilise the *Twin Life Study* and the *National Educational Panel Study* applying quasi-causal empirical designs. The findings challenge the liberal conception of merit as the sum of ability plus effort in evaluating equal opportunity.

Table of Contents

Acknowledgments

This thesis is nothing but a deep personal learning experience and a tiny contribution to sociological science. I am very grateful to those people who kept me on track and helped me to grow as a human being and researcher:

To Ildefonso, my academic father, and to Eva, for treating me as part of your family.

To my parents and my sister. Despite trouble, you always encouraged me in this long journey.

To Fabrizio, for your caring supervision in *Itagnolo*, boosting my intellectual horizons, and compensating for my self-boycotting instincts.

To Manolo, Echa and Javi, for initiating me in applied Sociology in the bars of Sevilla.

To my hometown friends in Dos Hermanas, as known as "los kikis", for funding with your taxes my so-called 4-year *Erasmus* in Florence.

To the Catalan crew, Pablo (N.P.S.), Toni, Albert and Guillem, for being my hosts and *Cicerones* as a first comer in Florence, and for some nights *muy cremas*.

To Pitu, Jota, Ana and Gonzalo, for being my Spanish family in San Marco, and not punching me when playing FIFA.

To the Trento squad, Marco, Lele, Giampi, Giulia, and Fra, for making me feel at home in Italy and for many SISEC to come as fun as the ones we shared in Napoli and Torino.

To my French mates, Camille, Thibault, Louis and Maeva, for avoiding the language of Shakespeare and learning together what at the beginning seemed to be a ridiculous Italian.

To Brunella, for bearing me during the lockdown and beyond, for being such a sweet human being, and teaching me the difference between *orecchiette* and *orecchini*.

To Mar, Irene, Ilze and Mariajo, for your friendship, grit and intellectual inspiration.

To the thesis Jury, Juho Härkönen, Leire Salazar and Jonas Radl, for your terrific feedback.

To Martina, Jennifer, and Gabriella, for being very human and efficient professionals.

To my classmates and friends in Barcelona, Adri, Jorge, Stephen, Michiel, Bram, Simon, Ansgar, Daniel, Khatia and Sibel, for the good times of youth and *Estrella* hangovers.

To Sandra and Pedro, for being so nice to me, a rookie just arrived to Barcelona, and cooking so good Argentinian pasta and *asado*.

To Tonino Carotone, for defining the leitmotiv of my life, especially in times of pandemia: *È un mondo difficile. Felicita' a momenti. E futuro incerto.*

To the 7-line bus drivers, for your fast-laps, always seeking the next frontier in speed and on-time services. Up and downhill, you teached me that life is a roller coaster.

To all the inspiring and lovely people that I came across in Sevilla, Barcelona, Tilburg, Berlin and Florence during my BA and postgraduate studies and I was too lazy to mention here. You are also in my thoughts, but I just wanted to get this thesis submitted ASAP. I am eager for the moment when someone will ask if there is a doctor in the room and I will be there, finally being a Doctor of Philosophy, useless but full of wisdom.

Chapter I

Introduction

Carlos J. Gil Hernández

Ascriptive forces find ways of expressing themselves as "achievement" (Halsey 1977:184).

1. Overview

During the post-World War II era, known as the *Golden Age of Capitalism*, social mobility odds were in a positive scenario. Keynesian policies fostering economic growth with progressive redistribution between social classes, occupational upgrading into highly-qualified professional and managerial sectors, and major welfare state reforms levelled the playing field (Esping-Andersen 2015). Particularly, this joint structural context of educational expansion and occupational upgrading brought about more *room at the top* and sustained upward social mobility, while reducing relative inequalities in the intergenerational persistency of socioeconomic status (Breen and Müller 2020; Breen and Luijkx 2004). Even when large inequalities in accessing post-compulsory education existed (Raftery and Hout 1993; Breen et al. 2009), education could be considered as a social lift for all social classes alike (Goldthorpe 2013). From this past situation, today it is still widely believed that boosting university enrolment is the best bet for raising social mobility rates in capitalist societies (Goldin and Katz 2008).

Unlike the *Golden Age*, contemporary post-industrial societies are undergoing economic slowdowns, welfare state retrenchments, and increasing income inequalities (Esping-Andersen 2007). Upward mobility rates levelled-off as the post-industrialisation process reached its peak in the late 1990s—consequently, the rate of growth of highly-skilled jobs slows down (Breen and Müller 2020). This context of stagnant structural change may lead to increasing competition for accessing top occupations in terms of income and status. Namely, if the size of the cake—*room at the top*—remains the same and opportunities are to be equalised, "as a mathematical necessity, downward mobility has to increase just as much as upward mobility (Goldthorpe 2016:105-107)." Under this state of affairs, social mobility becomes a *zero-sum game* in which the pressure to avoid intergenerational downward mobility or social demotion among middle and upper-class families may strengthen.

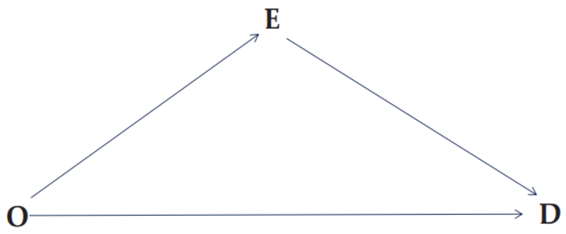


Figure 1. Origins (O) – Education (E) – Destinations (D) Triangle

In contrast to pre-industrial aristocratic societies where social positions were directly inherited by blood, nepotism or divine right, one of the main findings of social stratification research is that education is the main determinant of adult socioeconomic status (SES) in contemporary societies with mass schooling systems (Blau and Duncan 1967). However, building on the *status attainment model* and its Origins (O) - Education (E) - Destinations (D) triangle analytical framework (see Figure 1), social background is systematically associated with an individuals' socioeconomic position both indirectly (O-E*E-D path) and directly (O-D path) (Breen and Jonsson 2005). The indirect path operates via educational attainment (O-E*E-D), accounting for the largest share of the intergenerational association. Thus, at the same time, education might function as the main engine of social mobility or social reproduction (Hout and DiPrete 2006).

Nowadays, the most recurrent channel for upper classes to reproduce their status across generations and avoid falling down the social ladder is attaining high educational credentials. Likewise, for working classes, the best bet for upward mobility is getting ahead in education and access to college. Despite access to secondary education becoming more equal, and a driver of social fluidity in several European countries in the first half of the XX century (Breen et al. 2009; 2010), SES-inequalities in getting access to college education, the main social lift, have remained relatively unchanged (OECD 2018). These educational inequalities may even increase in a context of depressed economic growth and high income inequality, thus resulting in disparities in parental educational investments and expectations (Salazar, Cebolla-Boado and Radl 2019; Schneider, Hastings and LaBriola2018; Lucas 2017; Reardon 2011).

Two seminal and competing theoretical streams were outlined to explain why educational inequalities tend to persist over time: *cultural reproduction theories* (CRT) (Bourdieu and Passeron 1990), and *rational action theories* (RAT) (Breen and Goldthorpe 1997). Both theories have substantial explanatory power and have been systematically tested and (re)elaborated, but several black boxes remain to be unpacked to understand fully why inequalities are so "sticky" across generations. In this dissertation, I mainly draw from these sociological theories to explore how educational inequalities are produced, highlighting some weaknesses that I address by incorporating analytic elements from other disciplines.

Cultural reproduction theories emphasise the unequal stock and transmission of cultural capital across families in explaining academic achievement (Bourdieu and Passeron 1990). Teachers misconceive cultural capital as a signal of academic brilliance and, as a result, the educational system functions as an institution of reproduction of inequality by positively evaluating those children socialised in the dominant culture (Jaeger and Mollegaard 2017). Indeed, it is well-known that parental education, as a proxy for cultural resources, is more predictive of children's academic achievement than income (Francesconi and Heckman 2016; Erikson and Jonsson 1996).

Yet, CRT generally overstate the role of cultural capital in reproducing inequalities in the educational system (Jaeger 2011), and mechanisms through which cultural capital may lead to educational success are not well-identified (Jaeger and Breen 2016:1108). Last but not least, an additional drawback of CRT lies in its deterministic stance in which no room for individual choice is left. I argue that it may be the case that the construct of cultural capital is endogenous to the development of cognitive and non-cognitive skills rewarded in educational systems (Farkas 2003). In other words, the socialisation into a parental environment rich in cultural capital may not only affect academic success through knowledge of the highbrow culture and teachers' bias in judgements, but mainly by parenting strategies that facilitate the development of those very cognitive and non-cognitive skills that enhance learning and academic performance.

Rational action theories draw from the psychological concept of "loss aversion", as defined as an intrinsic cognitive bias in human beings due to evolutionary pressures (Kahneman 2011), to posit that upper-class families are generally risk-averse to social demotion (Boudon 1974). The upper class has more to lose in terms of status than working-class families due to *ceiling* and *floor effects*, respectively (Breen and Goldthorpe 1997). Thus, upper-class families are particularly interested in pushing their kids to take ambitious decisions at critical junctures of the educational system. While RAT recognise the role of cultural, psychological, and genetic factors in explaining SES-gaps in academic performance (Jackson 2013), they focus on parental and students' rational choice mechanisms that are bounded by their relative position in the class structure.

I argue that RAT do not suffice as an explanation due to (1) focusing on individual choice mechanisms and not considering teachers as relevant actors, as CRT do; and (2) modelling ability and choice as independent factors (Jackson 2013). Regarding the first point on RAT focus on individual choice, it is crucial to further explore how ability differentials are generated among families in the first place. Ability differentials account for a large share of total educational inequality, depending their contribution on the educational system design (Jackson and Erikson 2013). If educational policy is to compensate for early inequalities in skill formation, it is vital to grasp how and when they are generated. Therefore, this dissertation studies inequalities in academic ability, drawing from developmental psychology, behavioural genetics, and skill formation models.

Moreover, it is important to acknowledge the role of teachers and their judgement biases in shaping parental and students' expectations of success in the educational system (Spinath and Spinath 2005), as argued by CRT, educators are the main evaluators of merit or gatekeepers in the educational system. This dissertation tackles this limitation by (re)incorporating teachers as protagonist actors in the early stages of the status-attainment process.

Concerning the second point on the independence of ability and choice, most research following the bounded-rationality framework assumes that differences in educational transitions between working-class and upper-class children remain constant across the academic-ability distribution (Jackson 2013). In turn, *compensatory theories* argue that inequalities in accessing educational pathways leading to college are disproportional among low-skilled students (Bernardi and Triventi 2018), so that ability and choice might be interacting. The rationale behind these theories is that affluent families are particularly motivated to mobilise their extensive resources to prevent their kids from falling down the (educational) ladder due to risk aversion to social demotion (Goldthorpe 2007). This is particularly the case in the negative event of low scholastic ability when the risk of downward social mobility peaks. This dissertation contributes to the literature by extending the RAT framework with compensatory theories to explore whether and how advantaged kids tend to avoid downward mobility from early in life.

The dissertation bridges interdisciplinary literature from sociology, psychology, and economics to study how parental SES affects skill formation and educational achievement during childhood (age 5-11) in Germany and evaluates the implications of the empirical findings for social justice theories. The dissertation contributes to the literature on intergenerational educational inequality by (1) analysing the consequences of prenatal health shocks in the development of cognitive and non-cognitive skills; (2) examining the effects of cognitive and non-cognitive skills on transition rates to secondary academic education; and (3) assessing the stratification of these associations by parental SES.

Drawing from *compensatory advantage theories*, I study how negative shocks and traits— low birth weight (LBW) and low cognitive skills—may be less detrimental for high-SES kids from early stages of the status-achievement process. Particularly, I explore mechanisms such as parental educational investments and aspirations and teachers' bias in assessments. In analysing these issues, I engage in a complex normative debate about the definition of equal opportunity. Thus, I outline a normative framework ex-ante to evaluate equality of opportunity in education and test liberal theories of justice, which mainly conceptualise academic merit as the sum of natural ability plus effort.

Germany represents an ideal context to test liberal normative theories of equal opportunity due to its educational system that enforces tracking of children into academic or vocational pathways as early as at age 10. In this system of early tracking, SES-inequalities in accessing the academic track leading to college are thought to be mainly driven by SES-gaps in school readiness, and teachers are supposed to objectively assess students as a function of their ability and behaviour. Thus, tracking can be considered as an early starting gate to evaluate equal opportunity in education in which high-SES families may find it particularly difficult to compensate for low ability.

The dissertation consists of three individual empirical papers. In the first paper (Chapter II) co-authored with Marco Cozzani (20% contribution) and Fabrizio Bernardi (20%

contribution), we test whether high-SES families can compensate, through investments, for the negative effects of prenatal health shocks—LBW—on skill formation due to their large pool of economic and cultural resources. A socioeconomic gradient in the effect of BW on skill formation may contribute to the persistency of early SES-gaps over the life-course. We contribute to the literature by exploring two possible mechanisms that may account for the heterogeneous effect of BW by parental SES: (1) relative allocation of investments within families; and (2) absolute level of investments between families. We further contribute methodologically by exploiting random variation in twins' access to nutrients and oxygen *in utero* as a natural experiment to isolate random variation in BW.

In the second paper (Chapter III), published in *Sociology of Education*, I draw from the literature on educational inequality within families to test whether high–SES families compensate for low cognitive ability in the transition to secondary education. I use non-verbal intelligence quotient (IQ) tests as a proxy for natural ability and apply a quasi-causal twin design to assess whether compensatory mechanisms for low ability also work within families. I contribute to the literature by looking for the first time at the heterogeneity of the effect of IQ on track choice across parental SES and the absolute ability distribution within families.

In the third and last paper (Chapter IV), I provide novel findings on the interplay between cognitive and non-cognitive skills in predicting educational outcomes. It has long been argued that non-cognitive traits such as perseverance and motivation might outplay cognitive ability in explaining status-attainment. Thus, I test for the first time whether high-SES students with low cognitive skills have larger returns to non-cognitive skills than low-SES peers in the transition to academic secondary education. I further contribute to the literature by exploring mechanisms accounting for the compensatory hypothesis, such as teachers' bias and parental aspirations.

To carry out the empirical analyses in Chapters II-IV, I draw data from the register-based panel *Twin Life Study* (Hahn et al. 2016) and the *National Educational Panel Study* (NEPS) (Blossfeld, Rossbach and Maurice 2011), applying advanced quantitative methods and quasi-causal research designs to minimise unobserved confounding.

Finally, in Chapter V I outline a set of conclusions summarising the research questions, contributions, case study and empirical findings, while also discussing the implications for normative theories of equality of opportunity, policy interventions, and sociological and economic theories on intergenerational educational inequality.

The remainder of this introductory chapter is organised as follows. First, drawing from distributive justice theories, I delve into the role of families and skills in the delimitation of ascriptive and achieved factors to evaluate the concept of equality opportunity in education. Second, I elaborate on the joint role of cognitive and non-cognitive skills in educational attainment and their environmental and genetic sources of variation. Third, I provide a review of the state of the art in social stratification research, highlighting its main caveats and offering avenues for new research by drawing from neighbouring scientific disciplines. I focus on the role of families as the main social institution contributing to the reproduction of educational outcomes, though being constrained by the structural level of economic inequality and welfare policies in a given society. This review appraises an integrative theoretical and methodological framework by including the accumulated insights in the fields of sociological research in social stratification, developmental and personality psychology, behavioural genetics, epidemiology, and skill formation models in economics. Fourth, I comment on the school system as the second social institution shaping inequality of educational opportunities, focusing on the particularities of the German system of early tracking as a case study. Fifth, I explain the methodological setting of the dissertation and some related challenges. Sixth, I provide an overview of the thesis by summarising the research questions, methods, and findings of the empirical chapters.

2. Normative Framework

2.1. Equal Opportunity in Social Stratification Research

Social stratification is one of the most prolific fields in sociology. In an attempt to evaluate the level of equality of opportunity in industrial societies (Swift 2004), a vast amount of research has thoroughly studied the association between parental socioeconomic background and children's socioeconomic attainment, as well as its underlying mechanisms (Breen and Jonson 2005; Torche 2015). Despite this laudable endeavour (Hout and DiPrete 2006), normative considerations in the study of social stratification and mobility have not been very well integrated due to its predominant empiricist flavour. As put by Adkins and Guo (2008:237), "the differentiation of status determinants into social background characteristics and personal merits or abilities has long been a standard, if somewhat under theorized, feature of status attainment research."

In this vacuum, functionalist theories and status-attainment models (Parsons 1951; Blau and Duncan 1967) somewhat captured the normative debate around the conceptualisation and testing of equal opportunity drawing from liberal perspectives. Liberal theories argued how industrialisation and technological development demand the secular prevalence of merit-based selection in educational systems and labour markets (Bell 1972; Treiman 1970). Since the onset of the XIX century, compulsory schooling laws were progressively implemented (Rausher 2016), formally granting a minimum level of education for all social classes alike, that historically had been the privilege of the elite. Thus, according to liberal theories of industrialism, as far as educational systems warranted equality of opportunity and its selection criteria was based on demonstrated meritocratic criteria—academic ability as defined by the sum of IQ plus effort—the legitimation of the modern stratification system would be safeguarded (Parsons 1951).

By the same token, in the hiring process, employers would increasingly rely on educational credentials as main signalling instruments of ability and potential productivity. As a corollary, in this *education-based meritocracy* (Goldthorpe and Jackson 2008), *achieved factors* would gain weight with respect to *ascribed factors*, causing the E-D association to strengthen, and the O-E and O-D associations to vanish (see Figure 1 above). In the extreme of this liberal normative spectrum, Michael Young (1958) coined the term meritocracy in his dystopian satiric fiction about the British society after the implementation of the 1944 Education Act, which established psychometric-based tracking in the educational system at age 10—the Eleven-Plus Exam. Young's cautionary tale about the risks for social justice of applying the formula *merit = IQ + effort* to funnel individuals in the stratification system seems to be loosely applied in empirical research and disregarded in contemporary capitalist democracies. Meritocracy is a pivotal concept invoked by conservatives, liberals and social democrats alike (Wheem 2001), and its popular belief and support have not vanished in times of rising wealth and income inequalities (Piketty 2020; Mijs 2019).

In this dissertation, I argue that a crucial point against this normative framework is the inadequacy of liberal theories and status-attainment models to evaluate the concept of equality of opportunity due to its misleading interpretation of IQ, effort, and education as meritocratic or achieved factors (Mijs 2016). Those parameters might more likely represent ascription due to inequalities in environmental input and genetic transmission (Nielsen 2006:196, 2016).

2.2. The Race on the Playing Field: Ascription or Achievement?

Formal or legal equality of opportunity is one of the keystones under which contemporary democratic societies are built upon. From the late XVIII century, liberal revolutions brought about the "career open to talent", with its corresponding bureaucratisation process in the civil and military positions that removed aristocratic and guilds' privileges (Boli, Ramirez and Meyer 1985; Hobsbwam 1996). Meanwhile, the counterrevolutionaries praised the ancient regime by endorsing privilege and hierarchy as legitimate due to their functional and traditional values. Regarding the *Declaration of the Rights of Man and Citizens* of 1789, Hobsbawm (1996:59) writes that "men were equal before the law and careers were equally open to talent; but if the race started without handicaps, it was equally assumed that the runners would not finish together."

The race has been a recurring metaphor to illustrate the concept of equality of opportunity or the process by which individuals are distributed and legitimated among the ranks of the social ladder. But when does the race start? What attributes can be considered as handicaps against running in equal conditions? When can it be said that there is equal opportunity to legitimate later inequality of outcomes? Social stratification research differentiates between two ideal types to evaluate the fairness of the race: ascription and achievement/merit. Ascription is usually equated to attributes beyond individuals' control that depend on the natural and social lottery, such as gender, race, productivity-enhancing genetic endowments, and parental socioeconomic status that transmit abilities, aspirations, preferences, and cultural, social and economic resources. Achievement, or merit, is usually related to those factors associated with later socioeconomic attainment that are to a certain extent under the individuals' control, such as self-cultivation of physical, behavioural or psychological traits rewarded in the educational system and labour market, preferences, decisions or choices, and effort.

The boundaries between ascribed and achieved characteristics to evaluate the fairness of the race are permeable and far from fixed. Even though sociology on social stratification and political philosophy on distributive justice have developed valuable instruments to define and test the degree of equal opportunity in a given society, there is no consensus on what factors lie on each side of the ascription-achievement spectrum (Mijs 2016). Depending on the normative and moral standpoint, certain ascribed characteristics can be considered more or less fair in conditioning future socioeconomic outcomes and, correspondingly, more or less subjected to political intervention (Dardanoni et al. 2006). When social scientists study inequality they are implicitly applying and interpreting moral definitions of social justice. Thus, I aim at shedding light on this debate by clearly delimiting a definition of equal opportunity and identifying ascribed and achieved factors to test if the social contest is a fair or rigged one.

2.3. Theories of Equal Opportunity: Liberal and Luck Egalitarianism

On the one hand, conventional distributive justice revolves around the egalitarian liberal theory of John Rawls (1999:63), which can be summarised in the following way by how he defines equality of opportunity: "Assuming there is a distribution of natural assets, those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system." That is to say, that advantages resulting from circumstances of birth should not influence life prospects. Also, it implies that individuals can benefit from those attributes (e.g., academic ability) that they own unequally as a matter of luck due to nature. In particular, this approach stands against considering constitutive luck—individual's genetic, personality, and identity components—or the natural lottery of genetic endowments as unfair or subjected to compensation policy due to arguments about aggregated economic efficiency and self-ownership of genetic endowments (Swift 2005:263).

In the context of justice and equal opportunity in education, Brighouse and Swift (2014) draw from the Rawlsian liberal framework to define their *meritocratic conception* of fairness in education as follows: "An individual's prospects for educational achievement may be a function of that individual's talent and effort, but they should not be influenced by her social class background (Brighouse and Swift 2014:15)." Similarly, Swift (2003:24) considers that, if equal opportunity and meritocracy in education are to be achieved, "people with the same level of merit—IQ plus effort—should have the same chance of success."

On the other hand, the *radical* perspective is mainly represented by the branch of *luck egalitarianism*, which can be summed up as follows: "all inequalities due to differential luck are unjust and give justice grounds for equalization, while those inequalities resulting from responsible choices are just (Swift 2005)." Indeed, according to Roemer (1998, 2012), the aforementioned trade-off between ascription and achievement can be more clearly thought as circumstances, and effort/decisions. Circumstances can be defined as the features of the individuals' environments that influence their achievement, and for which neither policymakers nor the society would hold individuals accountable. Roemer (1998, 2012) distinguishes four channels through which circumstances exert an influence on (income) opportunities across generations, ordered by the consensus on the degree of individual accountability (from more to less consensus, from less to more individual responsibility):

> C.1. Parents affect the chances of their children through provision of social connections.
> C.2. Parents affect the chances of their children through formation of beliefs and skills in children through family culture and investment.
> C.3. Parents affect the chances of their children through genetic transmission of native ability.
> C.4. Parents affect the chances of their children through the instillation of preferences and aspirations in children (private sphere, family).

Depending on which of these channels are considered as circumstances or decisions, different notions of equality of opportunity may emerge (Dardanoni et al. 2008:60). With respect to the first (e.g., nepotism) and second channels, most ethical observers would agree on labelling them as circumstances. The second channel is recognised in the US legislative records as a component of the legal conception of equality of opportunity (Jencks and Tach 2006): "Every child should have an equal chance to develop the traits that employers value." As Jencks and Tach (2006) point out, this statement implicitly refers to *equal educational opportunity*; however, there is no consensus about what it substantially means. Thus, they suggest to better put it as *equal developmental opportunity*, so that "all children should have the same opportunity to develop their innate talents."

With respect to the third and fourth channels, few (liberal) ethical observers would agree on not holding children responsible for their innate abilities and/or preferences. Thus, the more the intergenerational association between parents and children takes place through genetic channels, the greater the gap between conventional and radical distributive justice streams (Swift 2005:265). In other words, while the conventional approach would regard a society in which the allocation of socioeconomic attainment were 100% dependent on genetic differences in productivity-enhancing abilities as the realisation of equality of opportunity, the radical approach would consider this state of affairs as unfair inequality and subject to political intervention.

2.4. Towards an Evaluation of Equal Opportunity

Two main limitations of conventional and radical normative streams should be highlighted to get closer to a critical interpretation and testing of the concept of equal opportunity. First, both liberal and radical theories draw a sharp line between effort/decisions (radical), and intrinsic ability (liberal) as indicators of merit, versus circumstances of birth as unfair forces. This line is substantially thinner given that effort and decisions are not independent of social environments or biology, but considerably constrained by them (Spinath 2005; Sapolsky 2017). Likewise, liberal theories explicitly consider cognitive ability as an innate natural talent and central indicator of merit in addition to effort. However, cognitive ability is not only biologically determined as its development is also conditional on environmental input. Hence, measuring its innate component, net of inequalities in developmental opportunities, to evaluate equal opportunity is not technically possible at the moment of writing these lines (Conley and Fletcher 2017). Therefore, the liberal concept of merit in education is misleading (Fishkin 2014:57-59). By comparing the strength of birth circumstances—parental SES—on students' educational attainment at the same level of merit, "natural talent" (e.g., non-verbal IQ) and effort, we underestimate the role of previous inequalities in opportunities for skill development (Bukodi, Erikson and Goldthorpe 2014).

Second, both conventional and radical distributive justice streams are starting-gate theories: let's provide fair life chances for everyone by equalising opportunity, and then the argument follows, let's define a starting gate from which we will all participate in a fair contest. In other words, they suppose that there is an initial scenario of equalisation of opportunities from which we can safely evaluate the fairness of the race by applying meritocratic criteria at key moments of decision and selection of the status-attainment process.

Some authors contend instead that this framing is inadequate since a fair evaluation of equal opportunity from an arbitrary starting gate (e.g., endogeneity with previous advantages) and its realisation (practicality) are virtually impossible (Mijs 2016). Even if it were possible, they argue that it would not be a desirable goal (Fishkin 2014:65-74). Testing and materialising the praiseworthy ideal of equalising opportunity would imply controversial measures such as abolishing the family as a social institution, or applying genetic screening before birth (Huxley 1932; Young 1958).[1] Where do we establish the starting gate to evaluate merit while not reflecting previous inequalities in the development of this very merit? Marshal, Swift and Roberts (1997:7) note that "liberals tend to endorse the value of equality of opportunity, with inequalities of outcome deemed legitimate if they reflect differences in merit, but cannot agree about the conditions that are necessary to ensure that kind of equality of opportunity or about what attributes are meritorious."

Families will always influence their children's genetic endowments, personality and learning opportunities. Thus, the family prevents the realisation of equality of opportunity through constitutive partiality (i.e., intimate, loving, familial relationship) and illegitimate favouritism (i.e., nepotism, wealth) (Swift 2005:260). The normative and methodological challenge for the conventional approach lies on differentiating between (legitimate) inequalities of outcomes in the parental generation and (illegitimate) inequalities of opportunities in the children's generation. In turn, luck egalitarians would not necessarily seek to reduce the net effect of the family if this entailed increasing the weight of another morally arbitrary trait (e.g. innate ability): "Making family background less important means making merit a more important determinant of people's position in the distribution of advantage. To the extent that the distribution of merit is itself a matter of luck [in the genetic lottery], this is simply replacing one kind of injustice with another (Swift 2005:266-268)."

Given the implausible means that the luck egalitarian's measurement—distinguishing between responsible choices and brute luck in circumstances of birth and natural talents—and realisation of equality of opportunity would entail, I consider the liberal approach more practical to delimit and test the concept of equal opportunity. As argued in conventional liberal theories, "given problems in identifying the relative contributions of different factors to people's marketable abilities, we have to regard what the market will pay for those abilities as

[1] The American film *Gattaca* (Niccol, 1997), based on Aldous Huxley's *Brave New World* (1932), provides one of the best illustrations of a dystopian society in which the social hierarchy (valid and invalids) is determined by DNA tests: "It didn't matter how much I lied on my resume. My real resume was in my cells. Why should anybody invest all that money to train me when there are a thousand other applicants with a far cleaner profile. […] No matter how much I trained or how much I studied, the best test score in the world wasn't going to matter unless I had the blood test to go with it. Of course, it's illegal to discriminate. 'Genoism' it's called. But no one takes the law seriously."

the best feasible indicator of how much they do indeed deserve (Marshall et al. 1998:178-179)." Alternatively, in the case of the educational system, we shall regard what schools and teachers will mark as academic ability as an indicator of how much merit[2] students display. However, it should be clear that, in so far as families will always influence the constitution of their children through genetic transmission and nurturing of cognitive abilities, personality, and preferences rewarded in the educational system and labour market, we cannot say that people can deserve class advantages or disadvantages on the basis of this random allocation and exercise of attributes.

To challenge the liberal conception of equal opportunity and merit in education, I firstly evaluate if children from different social backgrounds have the same chances of developing those very abilities considered as main indicators of merit in the educational system. In the second chapter, I assess how inequalities start to gestate in the womb by assessing the impact of a random prenatal health shock—twins' differences in access to nutrients and oxygen *in utero*, affecting advantaged and disadvantaged families alike—on cognitive and non-cognitive skills' formation. Particularly, I evaluate if BW, an indicator of child perinatal health, developmental potential, or natural assets in the Rawlsian vernacular, has a differential long-term effect on children's developmental opportunities by socioeconomic circumstances of birth. If so, it would be an illustration of how natural assets or endowments interact with social environments in shaping unequal opportunities to develop academic merit from the starting gate of life.

Secondly, in the third and fourth chapters, I evaluate if, after accounting for individual differences in IQ and effort—due to nature and nurture, wealthy students at the same level of (liberal) scholastic merit than less affluent classmates have more chances of transiting into academic paths leading to college in Germany. The German educational system sorts students into academic or vocational tracks from age 10; thus, it represents an early starting gate in which formal selection criteria is based on academic merit after four years of public elementary education, where fair life chances are supposed to have been ensured. Thus, I explore how the main evaluators of merit in the educational system, i.e., school teachers, transform students' skills into grades, and whether they present any bias in their judgements as a function of students' ascribed characteristics at the same level of ability and effort. Finally, I also test if

[2] Defined as "The quality of deserving well, or of being entitled to reward or gratitude (Oxford Dictionary)."

educational inequalities are concentrated among cognitively weak students, compromising the validity of cognitive ability as an indicator of merit due to parental compensatory strategies (Bernardi 2014). In doing so, I will be able to evaluate the liberal definition of equal opportunity in education and its developmental, starting gate, and meritocratic components.

3. Skills, Genes and Achievement

To evaluate the concept of equal opportunity, I account for the abilities, skills and psychological traits most rewarded in educational systems and labour markets. It is well-known that educational attainment is one of the best predictors of later socioeconomic attainment. Thus, I am particularly interested in studying those skills that explain academic ability and performance from early childhood, namely cognitive and non-cognitive skills. These skills can be considered as indicators of academic merit according to liberal theories of social justice.

The conservative view on the abilities associated with later attainment is that intelligence is the main predictor of education, occupational class, and income, being largely genetically inherited and unchangeable (Hernstein and Murray 1994; Jensen 1969; Saunders 1996). Thus, educational interventions would be predetermined to fail (Heath et al. 1985). This vision is flawed and out-dated for two reasons (Rowe et al. 1999).

First, as recent research shows, personality or non-cognitive traits are at least as important as cognitive factors in explaining status-attainment (Bowles and Gintis 2000, 2002; Bowles et al. 2001). According to the *correspondence theory* by Bowles and Gintis (1976, 2002), similar cognitive and non-cognitive traits are rewarded in the educational system and labour markets. The educational system provides a socialisation process into the industrial discipline later demanded by employers for their workers—obedience to authority for lower classes, and creativity or imagination for upper classes. According to Bowles and Gintis (1976, 2000), personality traits are more important than cognitive abilities in explaining socioeconomic attainment and the persistency of inequality across generations.

Indeed, conscientiousness is by far the big-five personality trait most associated with grade point average (GPA) and educational attainment, over and above IQ (Almund et al. 2011). Interestingly, some studies further argue that this association between conscientiousness and

GPA is as large as the one found between IQ and educational achievement (Duckworth and Seligman 2005; Duckworth et al. 2012; Heckman and Kautz 2012:457). Furthermore, cognitive skills are generally captured with IQ tests, scores on achievement tests, or GPAs, but these measures are far from being perfectly correlated. Actually, IQ is the worst predictor of educational attainment among these cognitive measures (Borghans et al. 2016; Rindermann 2007).

The *High/Scope Perry Preschool Intervention Program* implemented in the 1960s targeted disadvantaged families by providing them with a high-quality and intensive curriculum for two years (Duncan and Magnuson 2011). While the treated group of students underwent a short-term boost in IQ scores, it rapidly disappeared, catching up the control group—known as the *fade-out effect* (Protzko 2015). However, these treated children enjoyed long-term benefits such as lower incarceration rates, less welfare-dependence, and higher educational attainment and income than the control group. This evidence suggests that other factors rather than cognitive abilities, such as attention control or inhibition of aggressive behaviour, may also be important in shaping future success in school and labour markets (Heckman and Kautz 2012).

Second, intelligence or IQ is not fixed at birth, but it is malleable[3] and dependent on environmental quality (Farah et al. 2008; Capron and Duyme 1989; Guo and Stearns 2002; Kendler et al. 2015; Ritchie and Tucker-Drob 2018; Tucker-Drob, Briley and Harden 2013). From Plato's Republic, one of the most controversial and ancient debates in philosophy and social sciences revolves around the relative contributions of nature and nurture in shaping human differences in behaviour. Thanks to the theory of evolution by Charles Darwin, the role of genetics entered into the scientific and political realms. Darwin's half-cousin, Francis

---

[3] Adoption studies provide an illustrative example and solid evidence on both the malleability and heritability of intelligence as a function of the (parental) environment. Imagine a quasi-interventional setting in which two monzygotic -twins who were born in deprived families, but, at 4 years old, were given up for adoption. One twin finds himself in an adoptive home of wealthy cultural professionals, while the other one is adopted by a working-class family that struggles to make ends meet. If we measure their cognitive abilities pre- and post-adoption, and the former twin showed a considerable advantage, the quasi-causal claim on the enriched rearing environment driving this difference could find reasonable support. As Duyme et al. (1999) and Kendler et al. (2015) have shown, it is actually the case, so we can claim that intelligence is not fixed at birth, but it is malleable to a certain extent depending on the environmental exposure. Indeed, in a landmark investigation drawing from the *Minnesota Twins Reared Apart Study* on how less similar reared-apart twins are compared to those reared-together, of several personality characteristics, the only trait for which sizeable differences were found was IQ (Bouchard et al., 1990). Though, these findings on the environmental malleability of psychological traits do not exclude the role of genetic influences among non-MZ twins. Most likely, a child with low genetic predisposition ("low-IQ" genotype as measured by GWAS; see Sniekers et al. 2017; Okbay et al. 2016) for developing cognitive abilities raised in a highly stimulating environment would not catch up a child with high genetic predisposition raised in the same enriched environment.

Galton, was a pioneer in developing and applying new statistical methods to disentangle the determinants of individuals' differences in behavioural traits, such as intelligence. For this enterprise, Galton (1875) devised the foundations of the twin method (Waller 2012), the *workhorse* of behaviour genetics research.[4]

The twin design relies on the comparison of the degree of similarity in phenotypic traits between individuals of different degrees of genetic relatedness: monozygotic (MZ) twins sharing 100% of their genome vs dizygotic (DZ) twins sharing 50%, on average (Bouchard et al. 1990).[5] The variance of a particular trait is decomposed into three linear components, known as the ACE model (Knopik et al. 2017):[6] genetic variance between and within families (A); environmental factors shared by twins in the same family that differ from one family to another (C); and non-shared environmental factors that differ between twins within the same family (E), such as twin-specific friends and teachers, or special parental treatment and reactions, plus measurement error (≈10%).[7] Under certain rigid assumptions[8], heritability represents the proportion of variation in a given trait that can be explained by genetic differences among individuals in a given population and time.[9]

[4] With the insights provided by the *Human Genome Project* and the increasingly cheaper gathering of DNA markers for large samples (biobanks), it is currently possible to measure DNA directly, instead of relying on the black-box of genes assumed by the classic twin-design (Conley and Fletcher 2017). The use of directly measured DNA led to the development of molecular genetics, using methods such as linkage analysis, candidate gene studies, genome-wide association studies (GWAS), and genome-wide complex trait analysis. The most promising avenues of research draw from the two latter methods. GWAS is a data-driven or atheoretical search of single nucleotide polymorphisms (SNPs; "individual segments of DNA that take only two values of the four (ATCG) that are possible" (Turkheimer and Harden 2014)) associated with personality traits among large samples of genetically-unrelated individuals. This approach has discovered the set of genetic variations associated wtih IQ and educational attainment (Okbay et al. 2016; Rieveld et al. 2013; Sniekers et al. 2017), for instance. However, GWAS present some methodological problems such as small effects and population stratification. Regarding the former, unlike rare diseases such as Huntington's disease, complex human behaviour is caused by thousands of genetic variations. Concerning the latter, the use of massive samples including individuals from different birth cohorts and national contexts (even though ethnic and ancestry homogeneity is usually controlled for) prevents finding the full variance in a given trait as previously established by twin studies ("missing heritability") (Turkheimer 2011). For instance, while twin-studies found that around 40% (Branigan et al. 2014) of the variance in educational attainment is explained by genetic differences, relying on polygenic scores derived from GWAS, Conley et al. (2015) just could account for around 3% of the genetic variance. The latest evidence suggest genetic effects do vary across populations and historical periods, so "that large homogenous datasets are required for behavioural phenotypes and that gene–environment interaction may be a central challenge for genetic discovery (Tropf et al. 2017)". Therefore, the use of the twin-design is still of great relevance in the *omics* era (Van Dongen et al. 2012).

[5] "Monozygotic (MZ) twins "develop from one embryo, which in the first few days of life splits into two embryos, each with the same genetic material." Fraternal or dizygotic (DZ) twins "develop from separately fertilized eggs. They are first-degree relatives, 50 percent genetically related like other siblings (Knopik et al. 2017:83-86)."

[6] The economists usually implement a fixed-effects model (within-MZ approach) to control for all potential sources of confounding (Jaeger and Mollegaard 2017), such as genetic and environmental endowments shared by MZ-twins in the same family (Kohler et al., 2011). Thus, any discordance in the dependent variable of interest between MZ-twins of the same family may be caused by the independent variable of interest. However, the main caveat of this method is the small variation in both the independent and dependent variables of interest.

[7] C is assumed to be the same for both MZ and DZ-twins given that they were born on the same day, so growing up under similar environmental circumstances.

[8] Equal shared-environments between MZ and DZ-twins, which is in general reasonably met (Conley et al. 2013); no assortative mating of parents (adjustable if the phenotypic correlation between parents is available in the dataset: r(DZ)=0.5+0.5*h2*rP) (Loehlin et al. 2009); and non-additive effects (genes-environment correlations or interactions, epistasis, and dominance deviations). Furthermore, it is also debated to what extent can be the patterns coming from the comparison between twins siblings be inferred to the general population, mainly made of full-siblings and singletons (especially so in a contemporary context of *lowest-low* fertility rates). In order to deal with this issue, the extended-family design was devised to compare the degree of similarity between twins and full-siblings within the same family (Hahn et al. 2016).

[9] $h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_C^2 + \sigma_E^2}$

A common misunderstanding about heritability estimates lies in interpreting that, for instance, because educational attainment is, on average, 40% heritable (Branigan et al. 2014), this estimate is measuring intergenerational genetic transmission. Instead, that a trait might genetically heritable means that both genetic variation between unrelated individuals across families, and random genetic variation between siblings born and raised in the same family, contribute to the heritability estimate. This tells us nothing about genetic transmission between parents and children, or about whether one can or cannot change a trait throughout political interventions.

For Galton, the high heritability of intelligence—IQ is 50% to 70% genetically heritable (Björklund et al. 2010)—was the proof of nature outweighing nurture in the forging of the genius. In a context of Victorian biological racism, he advocated for selective breeding, coining eugenics. According to Galton's determinism of nature, the high genetic heritability of intelligence across generations would prevent political interventions from being successful (Jensen 1969). Social Darwinist theorists such as Herbert Spencer, pseudo sciences such as phrenology, the institutionalisation of eugenics by the Nazi regime, the political use of genetic explanations of human differences to justify inequality, racism and prejudice, and several flawed investigations, have all contributed to the preponderance of nurture, cultural or environmental explanations of human differences from the end of the second world war (Sapolsky 2017). Following the blank slate notion, social scientists have been blind to the role of biological factors in shaping human differences in behavioural and socioeconomic outcomes[10] by emphasising the unique importance of culture and nurture over nature.

This Manichean understanding of nature and nurture as opposed poles resembles the above-discussed normative debate on ascription and achievement, or circumstances and effort/decisions. Following this dichotomous thinking, Nielsen (2006:196, 2016) suggests that, when we attempt to estimate the level of social ascription or evaluate the concept of equality of opportunity through measures of parental SES, we should partial out the role of genetic factors from the overall association. He suggests so because some interpretations of liberal theories of justice consider that a society where the only source of variation in status-attainment was genetic would be the best approximation to realise equality of opportunity (Diewald et al. 2015).

[10] "Genes do not exert a direct influence on educational attainment. Instead, genes associated with educational attainment may influence different biological factors that in turn, affect psychological characteristics that finally influence educational attainment (Conley et al. 2015)."

This bipolar understanding of the causes of complex individual human traits and resulting macro-level characteristics of the social system, such as social mobility rates or income inequality, is too simplistic. According to the first law of behaviour genetics: "All human behavioural traits are heritable (Turkheimer 2000:160)."[11] Thus, "A correlation between parents and children cannot be simply seen as "prima facie evidence for sociocultural causal mechanisms (Turkheimer 2000: 162)." For instance, the correlation between parental SES and children's intelligence stands at around 0.33 (Neisser et al. 1996:82); thus, it may not only be driven by environmental transmission. Likewise, though, genetic "heritability [of productivity-enhancing traits] cannot be seen as prima facie evidence of causal genetic mechanisms (Diewald et al. 2015)." Heritability estimates can be used as a starting descriptive point, but we need to go way beyond these to reach a substantive understanding of complex human behaviour and its determinants (Johnson et al. 2009).

Two important points should be crystal clear. First, all productivity-enhancing behavioural traits related to educational and socioeconomic attainment are genetically heritable to a certain extent (Krapohl et al. 2014).[12] Second, at the current state of the art, we cannot fully disentangle the limits, relative weights, and causal links between nurture and nature. This is so due to non-additive effects, which potentially violate the central linear assumption of the ACE model, such as correlations and interaction effects between genes and environments (Tucker-Drob and Harden 2012a; Conley 2016; Conley and Fletcher 2017). Without directly measuring the genotype of parents and children (Conley et al. 2015), or employing exogenous environmental shocks, we cannot rule out confounding between genes and environments (Fletcher and Conley 2013). Namely, we cannot know what the particular role of genetic transmission, random genetic variation, and environmental transmission is in the status-attainment process. What we can certainly know is that nature and nurture, as ascription and achievement, are tightly interwoven.

[11] The first law can be more accurately defined as: "The degree of similarity between two people on any trait is monotonically related to their degree of genetic relatedness (Turkheimer 2016:24)." This definition prevents vague causal claims and unifies classical quantitative genetic methods with recent developments in genome-wide association studies (GWAS) among genetically unrelated individuals (Okbay et al. 2016; Rietveld et al. 2013).

[12] The heritability estimates of educational attainment stand at around 40%, on average (see Barnigan et al. 2013). Some authors argue that around 2/5 of the intergenerational resemblance in earnings can be explained by genetic transmission (Björklund et al. 2005; Jencks and Tach 2006).

4. Theories, Limitations and Contributions

Inequality of educational opportunities is shaped by the interaction of two critical social institutions: families and schools. Two seminal and rival theoretical streams were erected to explain the persistency of educational inequalities across generations: *rational action theories* (RAT) (Boudon 1974; Erikson and Jonsson 1996; Breen and Goldthorpe 1997) and *cultural reproduction theories* (CRT) (Bourdieu and Passeron 1990). Even though both theories and further developments have great explanatory power and have been systematically tested and (re)elaborated (Sullivan 2001; Jaeger and Breen 2016; Barone et al. 2018), validating and refuting some of their foundational postulates, three main weaknesses that I address in this dissertation should be highlighted.

Firstly, neither of these theories pay enough attention to the mechanisms underlying the association between social background and academic ability. Secondly, RAT do not suffice as an explanation due to (1) focusing on individual choice mechanisms and not considering teachers as relevant actors; and (2) modelling ability and choice as independent factors. Thirdly, the vast majority of status-attainment research has evaluated inequalities between families by drawing a random individual from each family (Sieben and de Graaf 2003). Thus, it is assumed that siblings are equally influenced by the resources and processes of the family, but, as we will see below, this is not necessarily the case. In the remainder of this chapter, I will review the main theories on SES-inequalities in skill formation and choice in educational transitions drawing from interdisciplinary literature and highlighting their limitations and the main contributions of this dissertation.

4.1. Skill Formation and Environmental Mechanisms

Most sociological research on educational inequalities has not paid enough attention to the early childhood period until recently (Skopek et al. 2016). As ongoing research in the fields of cognitive neuroscience (Hackman, Farah and Meaney 2010; Nelson and Sheridan 2011) and developmental psychology (Farah et al. 2008), and skill-building models in economics (Knudsen et al. 2016; Cunha and Heckman 2007) illustrate, different early childhood socialisation experiences by parental SES have an accumulative effect on learning and skill formation. These unequal experiences, which gestate already in the womb (Conti et al. 2018), shape early SES-gaps in academic ability that are observable in pre-school (Lugo-Gil and Tamis-LeMonda 2008). Early gaps in skill formation will be difficult to bridge in the absence of high-quality educational interventions during childhood and early adolescence (Kulic et al. 2019; Cebolla, Radl and Salazar 2016), and they may have a limited impact (Skopek et al. 2016). Since early SES-gaps in skill formation are mainly shaped in childhood and remain relatively stable over time (Passaretta et al. 2020), it is crucial to understand how they emerge by taking an interdisciplinary approach accounting for the interplay between developmental and environmental processes.

According to theories of skill formation (Heckman 2007), the current stock of children's human capital is a dynamic function of previous skills, genetic endowments and parental investments. Namely, the level of skills at a certain stage has a direct effect on the level of skills at a subsequent stage, along with an indirect effect through the parental environment. Skills are multidimensional and embrace health, cognitive, and non-cognitive attributes (Francesconi and Heckman 2016). Parental investments are also multifaceted and depend on preferences and resources, encompassing parenting and schooling quality.

This multistage interdependent process of skill formation has three main theoretical features (Hernández-Alava and Popli 2017). First, "self-productivity" is the property of skills to cross-fertilise across developmental stages; or the well-known leitmotiv *skills beget skills.* For example, a high-level vocabulary at age 4 fosters the level of reading abilities at age 6 because the individual is able to learn faster and more efficiently. Second, "cross-effects" refers to the virtuous circle or feedback loops between different sorts of skills (i.e., good attention control eases the development of cognitive skills). Third, "dynamic complementarity" refers to the productivity of investments or interventions, and it implies that the level of skills at a certain stage of life increases the productivity of investments at following stages. For instance, those children with more initial ability would benefit more from the stimulation of their parents. *Dynamic complementary* also predicts the productivity of the investment to increase over the life cycle so that compensatory investments tend to lose effectiveness over time (Aizer and Cunha 2012). By the same token, *dynamic complementarity* suggests that low levels of parental investments early in life can have long-term effects that are difficult to mitigate later in life. These predictions are based on the concept of critical or sensitive periods of child development, which can be understood as windows of developmental opportunity for brain malleability and the central nervous system, among others (Knudsen et al. 2006).

Although theoretically appealing, there is insufficient direct empirical support for *dynamic complementary* and *self-productivity* up to now (Bailey et al. 2017; Aizer and Cunha 2012). Some scholars argue that the longitudinal correlation between early skills and psychological characteristics in early childhood with these same traits at much later ages may be seriously overestimated (Watts et al. 2018), so compromising any causal claim (Duncan et al. 2007). This is due to the potential bias of unmeasured persistent factors, such as "differentially stable general cognitive abilities, personality, and environmental affordances (Bailey et al. 2017:2)." As we saw above, the stability of general cognitive abilities and personality traits is shaped by genetic factors to a great extent, depending on their expression on the environmental quality.

Drawing from sociology, economics and developmental psychology, Farkas (2003) assembles a common theory of family resources and child-rearing or parenting practices that influence children's development of skills and habits related to later socioeconomic attainment: economic, social and cultural resources or capitals. It is crucial to theoretically conceptualise different specific mechanisms associated with parental resources to prevent data-driven analyses. I pay special attention to economic and cultural capitals, as there is enough evidence coming from sociology, economics and developmental psychology to consider these resources as key explanatory factors of child development, academic performance, and later attainment (McEwen and McEwen 2017). These dimensions of social background are highly inter-correlated, thus, given the difficulties of testing experimental manipulations of economic or cultural resources (Duncan and Magnuson 2012), the best we can do is to try to isolate mechanisms of parenting.

4.1.1. Parental Economic Resources

Building on the human capital theory (Becker 1964), economists emphasise the level of parental economic resources and the different investments of money and time that they allow. Wealthy families with higher economic capital can afford sustained financial and time investments to enhance the learning opportunities of their children (e.g., enriching educational toys, private tutors and schools, summer camps, extracurricular activities, bedtime reading). Conversely, low-income and, particularly, those families who cope with poverty and deprivation (McEwen and McEwen 2017) on a daily basis, tend to work extended hours, have unstable schedules, long commuting times, and so on. Thus, they are more prone to suffer from chronic (toxic) stress, marital conflict, and related psychological problems (i.e., frustration, anxiety, depression), so leading to fewer resources and time to invest, affecting parenting quality (i.e., less warmth, monitoring, stimulation, chaotic home) (Layte 2017). This situation of sustained economic harshness negatively affects children cognitive (i.e., neurocognitive functions and IQ; resilience to stressful events) (Hackman, Farah and Meaney 2010) and non-cognitive development (i.e., aggressive behaviour, self-control, emotional regulation) via parenting practices. This is especially the case in liberal welfare states such as the US, while in the European context of more comprehensive social policies, the effects of poverty on children's development are more moderate. This suggests that income is more important when it leads to actual poverty.

4.1.2. Parental Cultural Capital

Cultural Reproduction Theories argue that upper-class families dispose of cultural capital, and schools and teachers positively evaluate those children socialised in the dominant culture (Bourdieu and Passeron 1990)—the ones who know the "rules of the game" (Lareau 2015). Cultural capital is expressed in three dimensions (Jaeger and Breen 2016): (1) embodied through socialisation: *habitus*; (2) objectivised in material cultural resources: books, pieces of art, musical instruments; and (3) institutionalised or formal: certified educational credentials.

It is well-known that parental education[13] as a proxy for cultural resources is more important than income[14] as a proxy for economic resources and investment capacity in explaining children's academic performance (Erikson and Jonsson 1996). However, given the lack of clarity and precision (Goldthorpe 2007) in Bourdieu's writings (2012), there is no consensus on the specific mechanisms that mediate the transmission of cultural capital between parents and children, and the relationship between children embodied cultural capital and demonstrated academic performance. According to Jaeger and Breen (2016:1108), "Research has yet [...] to identify the specific mechanisms through which cultural capital may lead to educational success."

Previous research has examined the following dimensions in the transmission of cultural capital between parents and children: highbrow culture, reading habits (e.g., bedtime reading), educational material resources (e.g., books, table games), cultural communication (i.e., teaching them to be analytical, reasoning and argumentative), and extracurricular activities (e.g., theatre, conservatory, second-language lessons). But what is the parental motivation to pass on their stock of cultural capital? While Bourdieu (2012) would argue that parents

[13] Net effect, controlling for income.
[14] Net effect, controlling for education.

relatively unconsciously reproduce behaviours interiorised during socialisation, Jaeger and Breen (2016) add rational choice behavioural assumptions.

Middle-class parents are *utility maximisers* who transmit the maximum possible amount of cultural capital. Whereas working-class parents are less likely to value certain cultural skills and behaviours as useful for future academic performance, middle-class parents follow a strategy of concerted cultivation[15]. In other words, different class-based sub-cultures in styles of child-rearing and biased teachers' evaluations by misconceiving cultural capital as academic brilliance would explain the distribution of academic performance at school and consequent educational attainment (Jaeger and Mollegaard 2017). Jaeger (2011:295-296) notes that future research should identify the particular mechanisms by which cultural capital influences academic achievement, such as teachers' bias (Jaeger and Mollegaard 2017), and parental educational strategies.

Some authors further argue that the construct of cultural capital is endogenous to the development of cognitive abilities and non-cognitive traits rewarded in the educational systems (Farkas 2003). In other words, the socialisation into a parental environment rich in cultural capital may not substantially affect academic success through knowledge of the highbrow culture and teacher's bias, but mainly by parenting strategies that facilitate the development of those cognitive and non-cognitive traits associated with later academic skills or competences.

I argue that the interpretation of cultural capital as culturally shaped skills (e.g., academic skills of language[16], reading, and mathematics), habits (e.g., homework, organisation, participation, effort, discipline), and knowledge is rather more appropriate (Farkas 2003:545). Beyond the traditional conceptualisation of cultural capital as highbrow culture, "Parental assistance with more mundane skills, for example, reading, is more consequential for students' success (De Graaf et al. 2000) (Farkas 2003:545)."[17] Indeed, among infants at 18 months of

[15] Lareau (2003:238): "In these middle class families, parents actively fostered and assessed their children's talents, opinions, and skills. They scheduled their children for activities. They reasoned with them. They hovered over them and outside the home they did not hesitate to intervene on the children's behalf. They made a sustained and deliberate effort to stimulate children's development and to cultivate their cognitive and social skills."

[16] "Hart and Risley (1995) showed that the child of professional parents has heard 30 million words by the age of three, the child of working-class parents has heard 20 million words, and the vocabulary is much richer for the richer SES child (Nisbett et al., 2012:136)."

[17] "Of course, the parents' own cultural capital (school-related skills and habits) is central to the provision of such parental assistance (Lareau and Horvat 1999; Lareau 2000, 2001). Furthermore, the relative and absolute value of the skills and knowledge in question continue to be points of contention in this research area. Thus, low-income parents may have real skills at surviving on a low income and coping with their life situation near the bottom of the stratification system, while still lacking the school-related skills necessary to help their children succeed at school. Meanwhile, high-income parents may have beaux-arts skills and habits (such as attending and appreciating high-culture music and art) that are of little productive value yet allow their children to signal high cultural status to their teachers. Lying between these two extremes are the more basic literacy and mathematics skills and habits that are correlated with parental social class, transmitted from parents to children, valued by teachers, and lead to real productivity increases in the worlds of school and work (Farkas, 2003:546)."

age, differences by parental SES in language processing skills (i.e., phonological awareness) and vocabulary have been already found (Fernald, Marchman and Weisleder 2013). Also, as early as 6-14 months of age, differences by SES in the brain executive function of working memory and inhibition control (i.e., attention skills) were identified (Hackman, Farah and Meaney 2010:652). The challenge lies in finding the specific parenting mechanisms that might mediate the relationship between social background measures and children's cognitive and non-cognitive outcomes.

4.1.3. Parental Cognitive Stimulation

The field of developmental psychology has studied for decades those cognitive skills and behavioural traits related to academic performance, and how the parental environment shape them from early childhood throughout every stage of development. Three main features of parenting have been identified as enhancing children's development of language, cognition, and school readiness: sensitivity (i.e., emotional support), cognitive stimulation (i.e., learning-promoting activities and environment), and warmth (i.e., affection and respect) (Lugo-Gil and Tamis-LeMonda 2008:1066). While most behavioural genetic and cognitive neuroscience studies have narrowly conceptualised the parental environment via SES (Duncan and Magnuson, 2012), developmental psychologists have devised directly observed measures of the family environment with a high degree of reliability.

Home Observation for Measurement of the Environment (HOME) is usually implemented in the study of pre-schoolers development and their families. This inventory accounts for the amount of intellectual stimulation (i.e., talking to the child; maternal speech and vocabulary); the level of access to books, magazines, newspapers, and computers; learning activities outside the household (i.e., museums, visits to friends); and the degree of maternal warmth or emotional support (i.e., encouragement vs reprimands); among others (Nisbett et al. 2012:136). It was found that HOME varies substantially across parental social class; and that a 1 standard deviation difference in HOME scores was associated with a 9-point difference in IQ scores. This latter finding is compromised by potential genes-environment correlations. However, it is plausible that a significant fraction of these IQ differences "is due to the environments independent of the genes associated with them (Nisbett et al. 2012:136)." In fact, even after accounting for maternal IQ and quality of parental care, Farah et al. (2008) found that the effect of parental stimulation on low-SES children's cognitive ability, as measured by language skills, remains predictive. It is surprising how the psychological HOME construct resembles the sociological concept of cultural capital, and the lack of dialogue between these disciplines.

4.1.4. Parental Educational Strategies

As introduced above, those children coming from disadvantaged families show more attention problems than their advantaged counterparts (Duncan and Magnuson 2011:12). However, research on the stratification of this personality facet, its role on the reproduction of inequalities, and the specific environmental mechanisms that may mediate the association between parental SES, children's personality, and academic outcomes is scarce (Kaiser 2016:3). As put by Shanahan et al. (2014:2): "Psychologists and economists have documented connections between personality and attained status but have not considered parental SES as an exogenous factor, and sociologists have documented effects on parental SES on attainment status but have not considered personality as a potential mediator."

As I pointed out above, it has been argued that middle-class parents with high cultural capital follow an educational strategy of concerted cultivation for their children (i.e., reasoning and discussing, structured activities, supervision of homework), while working-class parents are more likely to follow a "natural growth" strategy, which generally involves less supervision and organised time. These different parental strategies may shape differences in children's skills of attention control and, more generally, in the conscientiousness trait variance.

Bodovski and Farkas (2008) showed that a construct of concerted cultivation was predictive "of the children's approaches to learning, involving task persistence and attentiveness" (Kaiser 2016:4), a measure that can be linked to conscientiousness and attention problems. Likewise, Kaiser and Diewald (2014a) found that the effect of parental SES on conscientiousness, as measured by focus, was partially mediated by indicators of parenting practices similar to the ones highlighted by Lareau (2003, 2015). Kaiser (2016) also pointed to the same findings after applying a longitudinal design among children aged from 2 to 6. He found a strong effect of SES on child focus, but a small mediation estimate for the so-called "child-centred" parental model prevalent among middle and upper classes. Kaiser (2016:18) emphasises the importance of future research contributions taking into account the role of personality (sub)traits across different social environments in explaining the reproduction of social inequalities by integrating the insights of psychology and sociology.

4.2. Families and Choice in Educational Transitions

Girard and Bastide (1963) were forerunners in showing how the persistent relationship between social background and educational attainment could be unravelled into two components. *Primary effects* denote the systematic association between parental socioeconomic background and children's academic ability. *Secondary effects* account for the advantage of upper-class children in transition rates to higher educational levels than their working-class counterparts after controlling for ability. *Rational action theories* draw from this decomposition to focus on secondary effects or choice over and above ability differentials (Boudon 1974).

Rational action theories understand inequality in a bounded-rational action framework in which the relative position in the class structure constrains the resources, costs, benefits and chances of success to get ahead in the educational system. Building upon the psychological concept of *loss aversion*, RAT argue that upper-class families are less risk-averse when taking educational decisions in order to avoid downward social mobility—they have more to lose in terms of status maintenance. By contrast, for working-class families, less ambitious educational outcomes would suffice (*floor effect*) to reproduce or improve their status.

Goldthorpe (2007) points to three plausible rationales as to why working class children with similar academic performance than more advantaged children would systematically follow less ambitious educational tracks, or be more prone to dropout: (1) relative risk aversion, so that in order to avoid downward social mobility or demotion, less ambitious educational outcomes would suffice (*floor effect*); (2) less available and stable economic resources to afford the direct (i.e., tuition fees), indirect (i.e., living costs) and opportunity costs (i.e., earnings) to keep on studying; and (3) lower actual and perceived[18] chances of success due to their poorer average academic performance, along with underestimated or conservative perceived benefits of education. Except for *deviant* cases such as Sweden (Erikson and Jonsson 1996; Meghir and Palme 2005; Goldthorpe 2007), educational differentials by social origins would remain fairly stable.

As experimental research has shown (Barone et al. 2018), one can argue that secondary effects might also be explained by class-based differences in perceptions and information on the educational system and labour market (Lareau 2015). For instance, while working-class

[18] According to Gambetta (1987), these perceptions may not be completely rationally-grounded. Working-class families do not behave entirely rationally, but sub-intentionally over adapt, so that class-related values or norms about education, or psychological mechanisms would prevent them from making rational decisions (Erikson and Jonsson, 1996:16).

families may overestimate the difficulty of succeeding in upper secondary school and university, affluent families would push their children as they are more familiar with higher education. By the same token, (manual) working-class parents may underestimate the long-term labour market returns on academic-oriented tracks by attaching more importance to applied-technical skills learnt in vocational training. Recent developments of RAT (Breen et al. 2014:266) argue that those students with "lower time discounting preferences—those who prefer high economic rewards in the future to low returns in the present—are more likely to opt for academic secondary education."

4.2.1. Compensatory Advantage Theories

Most research relying on the RAT framework assume differences in transition rates between working-class and upper-class children to remain constant across the academic performance distribution (Goldthorpe and Jackson 2008; Jackson 2013). In other words, SES-gaps in transitions rates would be of a same size at low, medium or high levels of performance. However, as shown by Bernardi (2014) and associates (Bernardi and Triventi 2018; Bernardi and Cebolla 2014), SES-inequalities in secondary effects tend to be concentrated among low-performing kids. Low-ability advantaged kids are disproportionally more likely to opt for academic secondary education or have less risk of dropping out than working-class kids. That is to say that upper-class families tend to compensate for bad or mediocre academic ability to avoid intergenerational downward mobility given their extensive pool of economic, cultural and social resources: "They will give up only when persuaded by a clearly demonstrated lack of ability or interest on the part of their children (Erikson and Jonsson 1996)."

Other theories, such as *resource substitution* and *signalling*, have outlined similar predictions as the *compensatory advantage hypothesis*. Regarding the former, socioeconomic resources and skills may be complements or substitutes when it comes to predicting status attainment (Damian et al. 2015). On the one hand, in line with the *resource substitution hypothesis*, low-SES students might overcome their background disadvantage by relying on strong personality or cognitive skills, while skills may be less predictive of status-attainment for high-SES students, who can compensate with greater resources (Liu 2019). On the other hand, consistent with skill formation models, the *Matthew effect hypothesis* draws from accumulative (dis)advantage theories to predict that the "rich get richer" (DiPrete and Eirich 2006), so that skills are the strongest predictors of attainment among high-SES families.

New developments of educational decision-making models drawing from signalling theories (Spence 1973; Goldthorpe 2014) also predict compensatory patterns (Holm, Hjorth-Trolle and Jaeger 2019). Holm et al. (2019) argue that signals about academic ability and incomplete information about future chances of success are key mechanisms shaping choice and inequalities in educational transitions. Signalling models aim at explaining how students form beliefs about their own academic ability and chances of success, and how they may change these beliefs as a function of new signals, such as grades or information on the difficulty of educational pathways. Thus, Holm and colleagues (2019) argue that high-SES students might be less responsive to signals about academic ability than low-SES students because the former have a stronger drive to avoid social demotion, and their parents have resources to compensate for low ability. Furthermore, they predict that, under imperfect information about actual prospects of success, "information shocks" about the actual difficulty of educational tracks might depress low-SES student's chances of staying in education.

Generally, more research is needed on the precise mechanisms accounting for how secondary effects work and their relative weight (Barone et al. 2018). More specifically, little is known about what mechanisms of compensation are at play in case of poor ability or negative events among students from advantaged families. Enrolment in high-quality preschool programs, healthcare investments, parental involvement in cultural activities (Nicoletti and Tonei 2020), parental help with schoolwork and homework, private schooling and tutoring (Huang 2020), enrolment in extracurricular activities, residential and school choice, and parental aspirations are among the possible compensatory mechanisms suggested by previous research (Erikson and Jonson 1996; Bernardi and Cebolla 2014). However, little evidence exists so far on direct parental behavioural responses to compensate for their children's low ability (Bernardi and Grätz 2015), and whether these compensatory strategies are really effective to prevent them from downward social mobility.

In the remainder of this chapter, I will elaborate on additional mechanisms that might account for the predictions of the compensatory advantage model and that I will also test in the dissertation: (i) teachers' bias in assessments due to students' SES, (ii) and within-family allocation of schooling and health investments.

4.2.2. Teachers' Bias: An Elephant in the Classroom?

How do teachers, the principal evaluators of merit in the educational system, transform children's skills into educational success? Up until now, RAT have disregarded the role of teachers as central agents in the educational decision-making process. I draw from CRT and signalling theories to praise their importance in shaping inequalities in academic performance and educational decisions.

Teachers are the main gatekeepers or evaluators of academic merit in the educational system—ability + effort as defined by liberal theories (Swift 2003). Teachers, as all human beings (Sapolsky 2017), are exposed to implicit (subtle) biases in their judgment and behaviour, such as unconscious attitudes, reactions, and stereotypes in their perceptions of students' abilities and potential (Alesina et al. 2018). Such assumptions (e.g., statistical discrimination) may lead to self-fulfilling prophecies impeding student growth (Spinath and Spinath 2005). Previous research supports the existence of teachers' bias in grading standards and tracking placement recommendations as a function of students' ascribed characteristics, namely, gender, ethnic and socioeconomic background (Triventi 2019; Geven et al. 2018). Most previous research has focused either on ethnic or gender discrimination, calling for further research on SES-based bias.

Low-SES families are less risk-averse to downward mobility (*floor effects*) and have less perceived chances of success in education than high-SES families. Hence, they may be especially sensitive to distorting biases in the signalling information that teachers' evaluations provide (Holm et al. 2019), likely pushing their educational expectations downwards (Spinath and Spinath 2005). This *distorting effect* may be reinforced when low-SES students are low-performers, around a pass or fail grade to grant access to academic itineraries leading to college, where information on potential success is particularly unclear (Bernardi and Cebolla 2014).

Teachers' bias in grading is generally measured as the residual effect resulting from the difference between GPA assigned by teachers and blindly-assessed, standardised test scores (e.g., PISA). Nonetheless, the correlation between GPA and test scores is far from perfect, standing at about 0.63 (Südkamp et al. 2012). Thus, it is crucial to control for students' non-cognitive skills when assessing teachers' bias as the difference between GPA and test scores. Otherwise, this measure could just be reflecting the fact that students tend to exert less effort in low-stakes testing settings. Indeed, this is a crucial limitation of most prior research that this dissertation addresses specifically.

In the case of bias by socioeconomic background, CRT argue that teachers positively evaluate those children socialised in the dominant culture of the upper classes, to which teachers themselves belong. However, causal evidence on mechanisms is still scarce. As pointed out above, "research has yet [...] to identify the specific mechanisms through which cultural capital may lead to educational success (Jaeger and Breen 2016:1108)." Experimental research has taken the first step forward by evidencing intrinsic bias in teachers' evaluations of ethnic minorities through implicit association tests[19] (Alesina et al. 2018). Nevertheless, direct evidence on intrinsic bias by SES is lacking, and it is not clear-cut what the relative weight of unconscious bias (out-group bias in cognition) and conscious bias (explicit racism or classism) is in shaping teachers' judgments.

The causal basis of the CRT, claiming that cultural capital does cause educational performance and latter attainment, is compromised by the potential correlation of unobserved factors with both (parental and children) cultural capital and academic success. As Jaeger (2011:282) highlights, those children who display high levels of cultural capital are also highly likely to be the very ones with high innate ability and motivation—the same applies to their parents. This leads to a considerable overestimation of the total effect of cultural capital. Thus, we are facing a methodological problem to identify the effect of teachers' bias as a function of students' cultural capital or parental SES.

In order to deal with these issues, Jaeger (2011) implemented a "double fixed-effect" strategy by which he was able to control for within-family (sibling data) and within-individual (repeated observations over time) heterogeneity[20], so partialling out the major sources of latent confounding. He found that the estimated effects of cultural capital were less than half in comparison to the baseline models, which did not take into account unobserved heterogeneity within families and individuals. These findings point to the joint importance of unobserved factors and cultural environments shaping teachers' evaluations and academic success (Jaeger and Breen 2016; Jaeger and Mollegaard 2017:131).

[19] https://implicit.harvard.edu/implicit/iatdetails.html

[20] Generally, siblings-models allow to control for the shared family environment (C), but do not allow full accountability for genetic heterogeneity (A), as full siblings just share 50% of their genome on average, and are born at different points in the parental life-course (i.e., different resources) (Kohler et al. 2010:7).

4.3. Within-Family Inequalities: Theoretical and Normative Implications

Most status-attainment research studies the inequalities between families by drawing a random individual from each family (Sieben and de Graaf 2003). In doing so, it is generally assumed that siblings resemble each other in ability and behavioural traits, achieving similar socioeconomic outcomes and that they are equally influenced by the resources and processes of the family. However, by drawing a single individual from different families, we cannot control for those factors that siblings share or not. Conley (2004; 2008a) showed that the sibling-correlation in attainment measures such as education, occupation or income are far from perfect, standing only at about 0.5 in the USA (Conley and Glauber 2008b).

There are several factors that vary between siblings within the family that may explain this finding, namely the following: mother's age, birth order, birth spacing, sibling-specific shocks (e.g., divorce, economic crisis), and family climate. Gratz (2018) examined some of these factors and found that birth order (second born) and spacing (more closely spaced births) have a negative effect on educational outcomes, especially for disadvantaged families. Also, genetic makeup—on average, full siblings share 50% of their genotype—can make a difference within families. Indeed, leading behavioural geneticists such as Paige Harden consider that genetics is more useful in explaining individual differences in behaviour within-families than between-families.

Siblings who differ in their endowments can be treated differently by their parents, so niche picking effects may enter into play by reinforcing or compensating small initial differences in endowments in the long run (Grätz and Torche 2016; Conley 2004). Conley (2008) builds on the economics literature on intra-household allocation of resources to theorise about different patterns of within-family inequality by parental SES (Behrman et al. 1982). Similar to the *compensatory advantage hypothesis*, he suggests that advantaged families are more likely to *compensate* for siblings' differences in endowments thanks to their reliance on a large pool of cultural and economic resources. This would allow lower ability siblings reach the same educational outcomes as their more gifted siblings. Conley (2008a) found that the correlation in attainment measures for siblings coming from disadvantaged families is lower (reinforcement) than for those coming from advantaged ones (compensation), weakening over the life-course for the former group. However, the literature on parental response to children's endowments offers mixed findings on these intra-family dynamics (Almond and Mazumder

2013). This dissertation will shed light on this debate by applying a twin-design and analysing within-family inequalities in investments and educational outcomes by parental SES.

The study of within-family inequality also comes with substantial normative implications, being largely understudied by political philosophers (Landes and Nielsen 2012). Generally, a high correlation of status between siblings is interpreted as a sign of ascription or societal rigidity (Conley 2008). This approach can be misleading, as a high correlation between siblings (i.e., a pair arriving at the professional class) coming from a disadvantaged family should be better thought of as a positive outcome in terms of upward mobility and opportunity. Despite the inequalities produced within-families being a private affair in which the state has a limited impact (Swift 2005), it is important to understand and measure them to see a more complete picture on the mechanisms shaping the intergenerational transmission of (dis)advantages.

5. Educational Systems Design and Case Study

Regarding the school system as the second social institution shaping inequality of educational opportunities, different formal and informal institutional models may attenuate or reinforce social inequalities generated within families, so probabilistically conditioning the chances of staying or leaving the educational system, choosing or being allocated to qualitatively advantaged options (Triventi et al. 2016).

While educational systems with early tracking do reinforce the magnitude of primary effects or academic performance on educational inequality, comprehensive systems underpin the role of secondary effects or decision given that pupils follow similar tracks during lower-secondary education. Hence, there is a trade-off between the relative size of primary and secondary effects: more "meritocratic selection" or leeway to parental choice. Tracking systems seem to lead to larger overall inequalities (Bol and Werfhorst 2013). However, cross-country research has not provided highly comparable estimates of intergenerational educational mobility (Brunello and Checchi 2007; Pfeffer 2008; Jackson and Jonsson 2013).

From the Coleman Report (Downey and Condron 2016), it is well known that the largest proportion of educational inequality is shaped within-schools, suggesting the limits of educational policy and the salience of families. However, from the Swedish and Finish comprehensive reforms (Erikson and Jonsson 1996; Meghir and Palme 2005; Pekkarinen et al. 2009; Pekkala Kerr et al. 2013), causal evidence shows the positive effect of erasing early tracking and postponing the age of compulsory schooling on educational equality, or rising intergenerational income mobility over time.

This equalisation was not just limited to Scandinavian countries implementing comprehensive systems. According to Breen et al. (2009, 2010), several Central European countries also experienced a process of educational inequality reduction among the cohorts born in the first half of the XX century, though one was limited to the primary-to-lower-secondary transition. Causal methods were not applied, but they related ex-post this change to the declining salience of primary and secondary effects via the development of welfare states, democratisation and expansion of the educational institutions, and declining direct costs of studying due to the reduction of the average family size and sustained economic growth.

Overall, these findings on equalising educational opportunity challenge the assumptions of CRT and RAT on persistent inequality over time. Particularly, according to Goldthorpe (2007) and Erikson and Jonsson (1996), the deviant Swedish experience—characterised by sustained social democratic policies underpinning declining income inequality between the social classes, educational reform reducing direct and indirect costs of schooling, and employment security for working-classes (Esping-Andersen 2015)—is the exception that confirms the rule of generalised persistent educational differentials among industrialised societies (Shavit and Blossfeld 1993). These factors may have reduced the salience of secondary effects or choice (Rudolphi 2013). Primary effects, or SES-gaps in academic performance, would be less malleable by social reform since cross-national educational inequality seems to vary as a function of secondary effects (Erikson and Jonsson 1996; Jackson and Jonsson 2013:330). In line with this understanding, Piketty (2000:447) argues that there is not much to be done to mitigate the persistent inequality of abilities if it is "primarily determined by childhood learning through interaction with the parents at a very early age, and if this nurturing process is associated with the personality and behaviour of the parents rather than with material wealth per se."

One can argue that, instead, given the observed educational equalisation over time in highly diverse institutional settings, "the prevailing view that class inequalities in educational attainment will decline only under exceptional circumstances must be reconsidered (Breen et al. 2009:1514)." According to Goldthorpe (2007), educational expansion and reform tend to reduce the constraints that ability differentials exert on choice, as the degree of selectivity (ability and resources) in educational transitions is weakened. Namely, any potential reduction in SES-gaps in ability may decrease the share of students that were not able to make the choice of continuing in education due to objective or formal insufficient academic ability (i.e., institutional criteria for grade retention), and/or subjective-perceived chances of success (i.e., extreme caution of working-classes). Given that choice is conditional on ability, early educational interventions or high-quality universalised pre-school education may boost those cognitive and non-cognitive abilities related to educational achievement (Barnett 2011; Diamond et al. 2007; Gamoran et al. 2012; Pekkala et al. 2013; Schindler 2015).

Alternatively, in the current context of rising income and wealth inequalities in post-industrial societies, SES-gaps in parental educational investments may grow apart (Reardon 2011; Schneider, Hastings and LaBriola, 2018), so that underperforming kids from low-SES families may lag behind. At the same time, low-ability high-SES students may manage to get ahead thanks to compensatory strategies in a context of high risk of downward mobility and competition to access the upper-classes (Lucas 2017; Bernardi 2014).

5.1. Case Study: the German School as a Bottleneck and a Starting Gate

This dissertation focuses on the German educational system. Germany presents one of the highest levels of SES inequality in academic ability, as measured by the variation explained by SES in test performances in PISA, as well as in rates of high educational attainment among country members of the the Organisation for Economic Co-operation and Development (OECD 2018). Even when Germany has considerably reduced its levels of inequality in educational performance during the last years after some OECD reports' warnings, SES inequalities in the attainment of higher education have remained relatively constant during the last decades. Some authors have identified the specificities of the German educational system as one of the main factors explaining these high levels of educational inequality in a comparative perspective (Bol and van de Werfhorst 2013). The German system applies early tracking at ages 10 or 12 (grades 4 or 6) into academic or vocational pathways, and some federal states enforce binding recommendations (Blossfeld et al. 2016a; Ashwill 1999).

Some authors consider the system of early tracking as a bottleneck that hinders upward mobility through college and reinforces early SES-gaps in academic skills (Fishkin 2014:146-147). As a result, Germany displays low levels of upward educational mobility and relative social mobility (OECD 2018). In a context of automation of technical jobs (e.g., obsolescence and redundancy of specialised technical skills) and growing income inequalities in the skill premium between the highly and lowly educated-groups, early track allocation into vocational or academic education might be consequential for status-attainment (Gabay-Egozi and Yaish, 2020), especially in highly-industrialised economies and dual educational systems such as Germany. Vocational education certainly provides larger short-term labour market returns than academic education in terms of unemployment rates and earnings (Shavit and Müller, 1998). In the long-run, previous findings show that academic education leads to higher status occupations (high managerial and professional jobs vs white-collar and blue-collar jobs) and earnings (see Biewen and Tapalaga, 2020 for Germany), and lower rates of unemployment than vocational education (Hanushek et al., 2017). However, few studies have applied a life course perspective to explore the long-term returns of secondary education tracks (Golsteyn and Stenberg, 2017; Korber and Oesch, 2019).

In normative terms, the German educational system of early tracking can be thought of as a starting gate in which formal selection criteria is based on academic merit after four years of public elementary education (Fishkin 2014:146-147), where fair life chances are supposed to have been ensured. Due to these particularities, the German setting is especially suitable for testing the compensatory hypothesis and evaluating normative theories on skills and merit. As teachers are supposed to recommend secondary schools on the basis of objective criteria such as academic performance and behaviour, high-SES parents may have less room to compensate if their kids are low performers at the first important crossroad to avoid downward social mobility. The German case is also particularly relevant for testing theories of skill-formation due to the fact that, in early tracking systems, early SES-gaps in academic ability may be especially important in reproducing the persistency of educational inequalities. In empirical papers 2 and 3 (chapters III and IV respectively) of the dissertation, I will explain in more detail the functioning of the German educational system drawing from previous qualitative research (Ashwill 1999).

6. Methodological Setting and Challenges

6.1. Identification Strategies for Causal Inference

We are experiencing a trend towards hyper-specialised research in sociology using identification strategies and experimental research designs as in the (new) economics or psychology fields. This causal or experimental turning point is positive in methodological terms, as I had thought that sociology was lagging behind when it comes to rigorous research design and endogeneity. However, this trend towards hyper-specialisation incurs the risk of identifying very specific effects and mechanisms that are not generalisable to different populations and are not relevant for policy-making beyond specific interventions, lacking historical and political context.

The risks of this experimental drift led by the group of *randomistas* were highlighted by some critics[21] (e.g., Prize Nobel winners: Joseph Stiglitz and James Heckman)[22] of the Nobel Prize awarded to experimental (e.g., randomised controlled trials) economists Esther Duflo, Abhijit Banerjee and Michael Kremer for their approach and their surgical policy recommendations in a world where poverty and socioeconomic inequalities have systemic roots. Also, cutting-edge experimental research by Raj Chetty and colleagues on the geography of intergenerational social mobility in the USA and related policy implications recalls this surgical context-free flavour (e.g., the *moving to opportunity* randomised social experiment in the 1990s). I hope that I managed to find a balance in the trade-off between causal identification and external validity in my dissertation by combining deep theoretical, normative and policy debates with rigorous empirical designs.

In the dissertation, I exploit long-term panel data and the twin design (Carlin 2005) to deal with causality. While the advantages of panel data and modelling are well-known for the sociological reader (e.g., minimising reverse causation) twin models are becoming increasingly popular as an identification strategy in social stratification research (Jaeger and Mollegaard 2017). Nature provides an experimental setting with the incidence of twins (Knopik et al. 2017). Twin models allow to control for more unobserved confounding than most previous research that uses between-family estimates and sibling fixed-effects (Jæger 2011). Twins are born into the same family on the same day and share at least 50% of their genetic makeup. Thus, twin-pairs discordant in exposure can be thought of as a natural counterfactual in which the co-twins can be used as their own control/experimental group (McGue et al. 2010).

I argue that an ideal test of the *compensatory advantage hypothesis* would compare siblings who differ in nothing but their (observable) academic potential. Hence, the main benefit of studying inequality dynamics within families is the possibility of controlling for a larger array of characteristics shared by siblings who live under the same roof—neighbourhood, school,

[21] *https://www.opendemocracy.net/en/oureconomy/impoverished-economics-unpacking-economics-nobel-prize/*
[22] *https://www.theguardian.com/global-development/2018/jul/16/buzzwords-crazes-broken-aid-system-poverty*

genes, and parental environment—than between-family models allow. By implementing twin fixed-effects, I can control for environmental and genetic factors that vary between families, and at least 50% of genetic differences among twins in the same family. The remaining 50% of genetic sources of variation among DZ-twins born and raised in the same family can be considered as a random phenomenon given that, in the process of reproduction, each sibling randomly gets 50% of their segregating alleles from each parent (Knopik et al. 2017). However, within-family variation in endowments or parental investments might not be randomly assigned due to twin-specific confounding factors (Turkheimer and Harden 2014), and there are some additional concerns about the external validity of the twin design. I will deal with these methodological issues thoroughly in the dissertation.

6.2. Mediation analysis

Mediation analysis is a powerful tool to disentangle mechanisms, and it is applied in various academic disciplines. In the last years, though, new methodological developments in causal inference methods have put into question the usefulness of classic mediation analysis techniques (e.g., SEM, path analysis, Baron and Kenny's method) decomposing total effects into indirect and direct effects to identify causal mechanisms (Pearl and Mackenzie 2018). Three critical assumptions are neither outlined nor met in most investigations carrying out mediation analysis with the aim to estimate unbiased direct and indirect effects: (1) no unobserved confounders affecting *treatment* (X) and outcome (Y); (2) no mediator (Z)-outcome (Y) confounders (e.g., collider bias); and (3) no X-Z interaction or moderation.

In the empirical chapters of the dissertation, moderation, instead of mediation, is the main focus of my analyses to test the compensatory hypothesis. I only carry out a mediation analysis to give a descriptive account of the role of skills in mediating the association between parental SES and track choice, and of parental investments in mediating the total association between BW and skills. As I do not apply causal mediation techniques, I am very careful not to infer causal conclusions from my analyses and I clearly outline the assumptions that I make to interpret estimations. Additionally, in Chapter IV, I carry out some robustness checks to test for unobserved mediator-outcome confounding (*Seemingly Unrelated Regressions*) and estimate controlled direct effects of parental SES by running parametric regressions that allow for *treatment*-mediator interactions (Acharya et al. 2016).

6.3. Nonlinearities

Throughout the dissertation, I opt for linear probability models (LPM) instead of logistic models to predict dichotomous outcomes (e.g., track choice), and opt for additive interaction effects (e.g., metric exposure—skills—and moderator instead of dummies) to test the compensatory hypothesis in educational transitions. According to recent research by Beck (2020), it seems that the logistic approach is equivalent to LPM. However, any deviation from the linearity assumptions of these latter models could bias results and yield spurious interaction effects.

Two reasons explain my choice of linear models. Firstly, in Chapter III, logistic models yield cluster-specific (within-families) odds-ratios that are unrealistically high and only exploit variation among a subsample of discordant twins in tracking and IQ that limits sample size and power unnecessarily. Secondly, in Chapter IV, I control for school fixed-effects and include two- and three-way interaction terms. Unlike in LPM, interactions in logistic models are conditional on independent variables, "and they may have different signs for different values of the independent variables, and their statistical significance cannot be tested with a simple t-test (Gomila, 2019:9)." Furthermore, due to missing information on clusters with no variation in the outcome, I followed the linear approach given that I control for around 300 fixed effects (schools) and around 80 out of 300 do not vary in tracking outcomes. Additionally, in Chapters III and IV I carry out robustness checks using logistic models (e.g., conditional fixed-effects logit model, and fixed-effects logit model) and nonparametric specifications of the moderators (terciles, binning and kernel) to test for nonlinearities (Hainmueller, Mummoloand Xu 2018; Knol and VanderWeele 2012) and the main results of the dissertation hold.

7. Thesis Overview

The dissertation consists of three individual empirical papers whose research questions, contributions, methods and main findings I summarise below.

7.1. Chapter II. Birth Weight and Skill Formation: Biological Destiny or Parental Response? (Co-authored with Marco Cozzani and Fabrizio Bernardi)

Birth weight is a key predictor of child development and socioeconomic attainment later in life. However, the consequences of BW for status-attainment are not biological destiny. Educational interventions targeted at disadvantaged families are successful in offsetting the negative consequences of prenatal health shocks. Similarly, parents respond with investments to children's birth endowments, influencing their later skill formation. This article tests whether high-SES families are able to neutralise/compensate for prenatal health shocks thanks to their large pool of economic and cultural resources. A socioeconomic gradient in the effect of BW on skill formation may contribute to the persistence of early SES-gaps over the life-course. We study two cohorts at ages 5 and 11, drawing from the German *Twin Life Study*. We implement twin fixed-effects models to estimate the causal effect of BW by exploiting random sources of variation in intrauterine growth between twins. Results show that lower-BW co-twins have worse academic performance and more behavioural problems than their heavier-BW co-twins. At age 5, we observe a causal effect of BW on academic performance and behavioural problems that fades away for high-SES children at age 11. We argue that this compensatory pattern at age 11 may be explained by high absolute levels of investments by high-SES families (e.g., cultural activities and warmth), but not by their relative allocation of investments within families. Thus, we argue that biology is not destiny because (enriched) social environments may offset the detrimental effect of BW on skill formation.

7.2. Chapter III. Do Low-IQ But Advantaged Kids Get Ahead? A Twin Study on Early Schooling Inequalities

This article bridges the literature on educational inequality between and within families to test whether high–SES families compensate for low cognitive ability in the transition to secondary education in Germany. The German educational system of early-ability tracking (at age 10) represents a stringent setting for the compensatory hypothesis. Overall, previous literature offers inconclusive findings. Previous research between families suffers from the misspecification of parental SES and ability, while most within-family research does not stratify the analysis by SES or the ability distribution. To address these issues, I draw from the *Twin Life Study* to implement a twin fixed-effects design that minimises unobserved confounding. I report two main findings. First, highly educated families do not compensate for twins' differences in cognitive ability at the bottom of the ability distribution. Second, holding parents' and children's cognitive ability constant, pupils from highly educated families are 27% more likely to attend the academic track. This result implies a wastage of academic potential for disadvantaged families, challenging the role of cognitive ability as the leading criterion of merit for liberal theories of equal opportunity. These findings point to the importance of other factors that vary between families with different resources and explain educational success, such as non-cognitive abilities, risk aversion to downward mobility, and teachers' bias.

7.3. Chapter IV. Does Hard Work Beat Talent? The (Unequal) Interplay between Cognitive and Non-Cognitive Skills

It has long been argued that non-cognitive traits such as perseverance and motivation might outplay cognitive ability in explaining status-attainment. Cognitive and non-cognitive skills are key predictors of educational success and indicators of merit for liberal theories of equal opportunity. Nevertheless, even when accounting for SES inequalities in skill formation, disadvantaged pupils are less likely to make it to college. According to *compensatory theories*, SES-inequalities in educational transitions are disproportionally found among low-performing students due to status maintenance drives. However, little is known about the mechanisms accounting for this pattern. As cognitive and non-cognitive skills may be *complements* or *substitutes* in predicting educational outcomes, I test whether high-SES students compensate for low cognitive skills by high non-cognitive skills in the transition to upper secondary schools. I further contribute to the literature by exploring mechanisms such as teachers' bias and parental aspirations. I draw from NEPS to investigate a cohort of German students from grades 1-to-5, when early tracking is enforced. To minimise selective attrition bias and confounding, I apply inverse probability weights and school fixed-effects. I report four findings: (1) high-SES students at the same level of skills than low-SES classmates are more likely to opt for the academic track; (2) this inequality is largest among low-skilled students; (3) high-SES students are better able to compensate for low cognitive skills by high non-cognitive skills; (4) teachers' bias in grading and track recommendations, along with (over)ambitious aspirations of high-SES families, partially account for results. These findings challenge the (liberal) conception of merit as the sum of ability plus effort in assessing equality of opportunity in education.

References

Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3), 512–529.

Adkins, D. E., and Guo, G. (2008). Societal development and the shifting influence of the genome on status attainment. *Research in Social Stratification and Mobility*. 26, 235–255.

Alesina, A, M Carlana, E La Ferrara and P Pinotti (2018). "Revealing stereotypes: Evidence from immigrants in schools", *NBER Working Paper* 25333.

Almlund, M., Duckworth, A., Heckman, J.J., Kautz, T., (2011). Personality psychology and economics. In: Hanushek, E.A., Machin, S., Wößmann, L. (Eds.), *Handbook of the Economics of Education*, vol. 4. Elsevier, Amsterdam, pp. 1–181.

Almond, D., and Mazumder, B. (2013) Fetal Origins and Parental Responses. *Annual Review of Economics*, 5: 37-56.

Ashwill, M. A. (1999). *The educational system in Germany: Case study findings.* Washington, DC: National Institute on Student Achievement, Curriculum, and Assessment (ERIC Document Reproduction Service No. ED430906).

Barnett, W. S. (2011). Effectiveness of Early Educational Intervention. *Science*, 333(6045), 975– 978.

Barone, Carlo, Moris Triventi, Giulia Assirelli, (2018). Explaining Social Inequalities in Access to University: A Test of Rational Choice Mechanisms in Italy. *European Sociological Review*, 34(5):554–569.

Bayley et al. (2017). Risky Business: Correlation and Causation in Longitudinal Studies of Skill Development. *American Psychologist*, 73(1):81–94.

Beck, N. (2020). Estimating Grouped Data Models with a Binary-Dependent Variable and Fixed Effects via a Logit versus a Linear Probability Model: The Impact of Dropped Units. Political Analysis, 28(1), 139-145. doi:10.1017/pan.2019.20

Becker, G.S. (1964). *Human Capital Theory*. New York: Columbia University Press.

Becker, G.S., and Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *J. Lab. Econ.*, 4(3, Part 2): S1–39.

Bell, Daniel (1972). On Meritocracy and Equality. *The Public Interest*, 29: 29-68.

Behrman, J.R., Pollak, R.A., and Taubman, P. (1982). "Parental Preferences and Provisions for Progeny." *Journal of Political Economy*, 90(1):52.

Bernardi, F. (2014). Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth. *Sociology of Education*, 87(2): 74–88.

Bernardi, F., and Cebolla, H. (2014). Previous School Results and Social Background: Compensation and Imperfect Information in Educational Transitions, *European Sociological Review*, 30(2):207-214.

Bernardi, F. and M. Grätz (2015). 'Making up for an unlucky month of birth in school: causal evidence on the compensatory advantage of family background in England', *Sociological Science*, 2, 235–251.

Bernardi, F., and Triventi, M. (2018). Compensatory advantage in educational transitions: Trivial or substantial? A simulated scenario analysis. *Acta Sociologica*, doi:10.1177/0001699318780950

Biewen, M., and neé Tapalaga, M. T. (2020). Early tracking, academic vs. vocational training, and the value of 'second-chance' options. *Labour Economics*, 66, 101900.

Björklund, A., Eriksson, K. H., and Jäntti, M. (2010). IQ and family background: Are associations strong or weak? *The BE Journal of Economic Analysis and Policy*, 10(1), 1–12.

Björklund, Anders, Markus Jäntti, and Gary Solon (2005). Influences of Nature and Nurture on Earnings Variation: Preliminary Results from a Study of Various Sibling Types in Sweden." Pp. 145-64. *Unequal Chances: Family Background and Economic Success.* Samuel Bowles, Herbert Gintis and Melissa Osborne Groves, editors. Russell Sage.

Blau, P.M. and Duncan, O.D. (1967). *The American Occupational Structure.* Wiley: New York.

Bodovski, K., and Farkas, G. (2008). "Concerted cultivation" and unequal achievement in elementary school. *Social Science Research*, 37(3):903-919.

Bol, T., and Werfhorst, H. G. Van De. (2013). Educational Systems and the Trade-Off between Labor Market Allocation and Equality of Educational Opportunity. *Comparative Education Review*, 57(2): 285–308.

Boli, J., Ramirez, F. O., and Meyer, J. W. (1985). Explaining the Origins and Expansion of Mass Education, *Comparative Education Review*, 29(2), 145–170.

Borghans, L., Golsteyn, B. H. H., Heckman, J. J., and Humphries, J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences*, *113*(47), 13354 LP – 13359.

Bouchard, T., Lykken, D., McGue, M., Segal, N., and Tellegen, A. (1990). Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart. *Science*, 250(4978), 223-228.

Boudon, R. (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*, New York: Wiley.

Bourdieu, P. (1986). The Forms of Capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York: Greenwood.

Bourdieu, P. (2012). *La distinción. Criterio y bases sociales del gusto.* Madrid: Taurus.

Bourdieu, P., and Passeron, J.C. (1990). *Reproduction in Education, Society and Culture.* Beverly Hills: SAGE.

Bowles, S. and H. Gintis (1976). *Schooling in Capitalist America.* Basic Books, New York.

Bowles, S, and H. Gintis (2002). The inheritance of inequality. *Journal of Economic Perpectives*, 16:3–30.

Bowles, S. and H. Gintis (2002). Schooling in Capitalist America Revisited. *Sociology of Education.* 75:1-18

Bowles, S., H. Gintis, M. Osbourne (2001). The determinants of earnings: a behavioral approach. *Journal of economic literature*, 39:1137-1176.

Branigan, A.R., Kenneth J. McCallum, Jeremy Freese (2013). Variation in the Heritability of Educational Attainment: An International Meta-Analysis, *Social Forces*, Volume 92, Issue 1, 1 September 2013, Pages 109–140, https://doi.org/10.1093/sf/sot076

Breen, R. and Goldthorpe, J. H. (1997). Explaining Educational Differentials: towards a Formal Rational Action Theory. *Rationality and Society*, 9(3): 275- 305.

Breen, R. and Jonsson, J.O. (2005). Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility, *Annual Review of Sociology*, 31(1): 223–243.

Breen R., Luijkx, R., Müller, W., and Pollak, R. (2009). Non-persistent Inequality in Educational Attainment: Evidence from Eight European Countries. *American Journal of Sociology*, 114(5):1475-1521.

Breen R., Luijkx, R., Müller, W., and Pollak, R. (2010). Long-term trends in educational inequality in Europe: class inequalities and gender differences. *European Sociological Review*, 26:31–48.

Breen, R., and Luijkx, R. (2004). Conclusions. In R. Breen (Ed.), *Social Mobility in Europe*, 383-410. Oxford: Oxford University Press.

Breen, R., and Müller, W. (2020). *Education and Intergenerational Social Mobility in Europe and the United States.* Stanford, California: Stanford University Press.

Breen, R., Werfhorst, H. G. Van De, and Jæger, M. M. (2014). Deciding under Doubt: A Theory of Risk Aversion, Time Discounting Preferences, and Educational Decision-making, *European Sociological Review*, 30(2): 258–270.

Brighouse, H. and Swift, A. (2014) 'The Place of Educational Equality in Educational Justice', in K. Meyer (ed.), *Education, Justice and the Human Good.* Abingdon: Routledge, pp. 14–33

Brunello, G., and Checchi, D. (2007). Does School Tracking Affect Equality of Opportunity? New International Evidence. *Economic Policy*, 22(52): 781-861.

Bukodi, E., Erikson, R., and Goldthorpe, J. H. (2014). The effects of social origins and cognitive ability on educational attainment: Evidence from Britain and Sweden. *Acta Sociologica*, 57(4), 1–18. doi:10.1177/0001699314543803

Capron, C., and Duyme, M. (1989). Assessment of effects of socio-economic status on IQ in a full cross-fostering study. *Nature*, 340(6234), 552–554.

Carlin, J.B. et al. (2005). Regression models for twin studies: a critical review. *International Journal of Epidemiology*, 34:1089-1099

Cebolla-Boado, H., Radl, J., and Salazar, L. (2016). Preschool education as the great equalizer? A cross-country study into the sources of inequality in reading competence. *Acta Sociologica*, 60(1):41-60

Conley, D. (2004). *The Pecking Order: Which Siblings Succeed and Why*. New York: Pantheon.

Conley, D. (2008a). Bringing Sibling Differences In: Enlarging Our Understanding of the Transmission of Advantage in Families. (2008). In Conley D. and Lareau A. (Eds.), *Social Class: How Does It Work?* (pp. 179-200). Russell Sage Foundation.

Conley, D. (2016). Socio-Genomic Research Using Genome-Wide Molecular Data. *Annual Review of Sociology*, 42: 275–99.

Conley, D., Domingue, B., Cesarini, D., Dawes, C. T., Rietveld, C. A., and Boardman, J. (2015). Is the effect of parental education on offspring biased or moderated by genotype? *Sociological Science*, 2: 82–105.

Conley, D. and Fletcher, J., (2017). *The Genome Factor: What the Social Genomics Revolution Reveals about Ourselves, Our History, and the Future*, Princeton University Press.

Conley, D., and Glauber, R. (2008b). All in the family?: Family composition, resources, and sibling similarity in socioeconomic status. *Research in Social Stratification and Mobility*, 26(4), 297–306. doi:https://doi.org/10.1016/j.rssm.2008.08.003

Conley, D., E. Rauscher, C. Dawes, P. K. Magnusson, and M. L. Siegal. (2013). Heritability and the Equal Environments Assumption: Evidence from Multiple Samples of Misclassified *Twins. Behavior Genetics*, 43(5):415–26. http://dx.doi.org/10.1007/ s10519-013-9602-1.

Conti, Gabriella, et al. 2018. "Beyond Birthweight: The Origins of Human Capital." *Institute for Fiscal Studies Working Paper*:18-30.

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

Damian, R. I., et al. (2015). Can personality traits and intelligence compensate for background disadvantage? Predicting status attainment in adulthood. *Journal of Personality and Social Psychology*, *109*(3), 473–489.

Dardanoni, V., Fields, G. S., Roemer, J. E. and Sanchez Puerta, M. L. (2006). How demanding should equality of opportunity be, and how much have we achieved? In S. Morgan, D. Grusky and G. S. Fields (Eds.), *Mobility and inequality: Frontiers of research from sociology and economics* (pp. 59-82). Palo Alto, CA: Stanford University Press.

Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007, November 30). Preschool program improves cognitive control. *Science*, 318, 1387–1388. doi:10.1126/science.1151148

Diewald, M., Baier, T., Schulz, W., and Schunck, R. (2015). Status attainment and social mobility: How can genetics contribute to an understanding of their causes? *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 67(S1), 371–395. doi:10.1007/s11577-015-0317-6

DiPrete, T. A., and Eirich, G. M. (2006). Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. *Annual Review of Sociology*, *32*(1), 271–297.

Downey, D. B., and Condron, D. J. (2016). 'Fifty Years since the Coleman Report: Rethinking the Relationship between Schools and Inequality. *Sociology of Education*, 89(3):207–20

Duckworth, A. L., Quinn, P. D., and Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card. *Journal of Educational Psychology*, 104, 439–451.

Duckworth, A. L., and Seligman, M. E. P. (2005). Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents. *Psychological Science*, 16(12), 939–944. doi:10.1111/j.1467-9280.2005.01641.x

Duncan. G. J., Dowsett, C. J., Claessens, A., Magnuson, K.. Huston. A. C, Klebanov, P Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43, 1428-1446. doi: 10.1037/0012-1649.43.6.1428

Duncan, G. J., and Magnuson, K. (2012). Socioeconomic status and cognitive functioning: Moving from correlation to causation. Wiley Interdisciplinary Reviews: Cognitive *Science*, 3(3), 377–386. doi:10.1002/wcs.1176

Duncan, G., and Magnuson, K. (2011). The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems. In: Duncan G., Duncan G., and Murnane R. (Eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances* (pp. 47-70). Russell Sage Foundation.

Duyme, M., Dumaret, A., and Tomkiewicz, S. (1999). How can we boost IQs of "dull" children? A late adoption study. *Proceedings of the National Academy of Sciences*, USA, 96, 8790–8794. doi:10.1073/pnas.96.15.8790

Erikson, R., and Jonsson, J. (Eds.). (1996). *Can education be equalized? The Swedish case in comparative perspective*. Boulder: Westview.

Esping-Andersen, G. (2007). Sociological Explanations of Changing Income Distributions. *American Behavioral Scientist*, 50(5): 639–658.

Esping-Andersen, G. (2015). Welfare regimes and social stratification. *Journal of European Social Policy*, 25(1): 124-134.

Farah MJ, Betancourt L, Shera DM, Savage JH, Giannetta JM, Brodsky N, et al. (2008). Environmental stimulation, parental nurturance and cognitive development in humans. *Developmental Science*, 11: 793-801

Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual Review of Sociology*, 29, 541–562. doi:10.1146/annurev.soc.29.010202.100023

Fernald, A., Marchman, V. A. and Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16: 234–248. doi:10.1111/desc.12019

Fishkin, J. (2014). *Bottlenecks: A New Theory of Equal Opportunity*. Oxford: Oxford University Press.

Fletcher, J. M. and D. Conley. (2013). The Challenge of Causal Inference in Gene–Environment Interaction Research: Leveraging Research Designs from the Social Sciences. *American Journal of Public Health*, 103(S1):S42–45.

Francesconi, M., and Heckman, J.J. (2016). “Child Development and Parental Investment: Introduction.” *The Economic Journal, 126*(596):F1–F27.

Gabay-Egozi, L., and Yaish, M. (2020). Short and long-term consequences of high-school tracks for earnings in Israel. *Acta Sociologica.* https://doi.org/10.1177/0001699320920919

Galindo-Rueda, F., and Vignoles, A. (2005). The Declining Relative Importance of Ability in Predicting Educational Attainment. *The Journal of Human Resources*, 40(2): 335-353.

Galton F. (1875). The history of twins, as a criterion of the relative powers of nature and nurture. Fraser’s Mag 1875; 12:566–76. Reprinted with revisions in Galton F. The history of twins, as a criterion of the relative powers of nature and nurture. *J Anthropol Inst* 1875;5:391–406 (Reprinted in *Int J Epidemiol* 2012;41:905–11).

Gambetta, D. (1987). *Were they Pushed or Did they Jump? Individual Decision Mechanisms in Education.* Cambridge: Cambridge University Press.

Gamoran, A.; Turley, R.N.L.; and Fiel, J. (2012) Evidence-based school interventions to reduce achievement inequality. In: Maholmes, V. and King, R.B. (Eds.), *The Oxford handbook of poverty and child development:* 372-384. Oxford: Oxford University Press.

Geven, S., Batruch, A., and van de Werfhorst, H. (2018). *Inequality in Teacher Judgements, Expectations and Track Recommendations: A Review Study.* Amsterdam: Universiteit van Amsterdam.

Girard, A., and Bastide, H. (1963). La stratification sociale et la démocratisation de l'enseignement. *Population*, 18(3): 435-472.

Goldin, C., and Katz, L.F. (2008). *The Race between Education and Technology.* Havard University Press: Cambridge, MA.

Goldthorpe, J. H. (2016). Social class mobility in modern Britain: changing structure, constant process. *Journal of the British Academy*, 4: 89–111.

Goldthorpe, J.H. (2014). The role of education in intergenerational social mobility: Problems from empirical research in sociology and some theoretical pointers from economics. *Rationality and Society*, 26(3): 265–289.

Goldthorpe, J. H. (2013). Understanding – and Misunderstanding – Social Mobility in Britain: The Entry of the Economists, the Confusion of Politicians and the Limits of Educational Policy. *Journal of Social Policy*, 42(3): 431–450.

Goldthorpe, J.H. (2007). *On sociology.* Stanford, California: Stanford University Press.

Goldthorpe, J.H., and Jackson, M. (2008). Education-Based Meritocracy: The Barriers to its Realisation. In A. Lareau and D. Conley (Eds.), Social Class: How Does it Work? New York: Russell Sage Foundation.

Golsteyn BHH and Stenberg A (2017) Earnings over the life course: General versus vocational education. Journal of Human Capital 11(2): 167–212.

Gomila, R. (2019, July 11). Logistic or linear? Estimating causal effects of treatments on binary outcomes using regression analysis. *https://doi.org/10.31234/osf.io/4gmbv*

Grätz, M. (2018). "Competition in the Family: Inequality between Siblings and the Intergenerational Transmission of Educational Advantage." *Sociological Science*, 5:246–69.

Gratz, M., and Torche, F. (2016). Compensation or Reinforcement? The Stratification of Parental Responses to Children's Early Ability. *Demography*, 53: 1883-1904.

Guo, G., and Stearns, E. (2002). The Social Influences on the Realization of Genetic Potential for Intellectual Development. *Social Forces*, 80:881-910.

Hackman, D. A., M. J. Farah, and M.J. Meaney (2010): Socioeconomic status and the brain: mechanistic insights from human and animal research, *Nature Reviews/Neuroscience*, Vol. 11 (September 2010): 651-659.

Hahn, E., Gottschling, J., Bleidorn, W., Kandler, C., Spengler, M., Kornadt, A., Schulz, W., et al. (2016). What Drives the Development of Social Inequality Over the Life Course? The German TwinLife Study. *Twin Research and Human Genetics*, 19(6), 659-672. doi:10.1017/thg.2016.76

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu (2018). "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*. https://doi.org/10.1017/pan.2018.46

Halsey, A. H. (1977). "Towards meritocracy? The case of Britain", en Karabel, J. y Halsey, A. H. (eds,), *Power and Ideology in Education*, Oxford University Press, Nueva York, pp. 173-186.

Hanushek, EA, Schwerdt, G, Woessmann, L, et al. (2017) General education, vocational education, and labor market outcomes over the life-cycle. *Journal of Human Resources*, *52*(1): 48–87.

Heath, Andrew C., K. Berg, Lindon Eaves, M.H. Solaas, L.A. Corey, J. Sundet, P. Magnus and W.E. Nance. (1985). Education Policy and the Heritability of Educational Attainment. *Nature* 314 (6013):734-36.

Heckman, J. J. 2007. "The economics, technology, and neuroscience of human capability formation." *Proceedings of the National Academy of Sciences of the United States of America, 104*(33):13250-13255.

Heckman, J. J., and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451–464. doi:https://doi.org/10.1016/j.labeco.2012.05.014

Hernández-Alava, M., and Popli, G. (2017). Children's Development and Parental Input: Evidence From the UK Millennium Cohort Study. *Demography*, *54*(2), 485–511.

Herrnstein, R. J., and Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

Hobsbawn, E. (1996). *The age of revolution. Europe 1789–1848.* London: Abacus.

Holm, A., Anders Hjorth-Trolle, and Mads Meier Jæger (2019). "Signals, Educational Decision-Making, and Inequality." *European Sociological Review*, 35(4):447–460

Hout, M. and DiPrete, T.A. (2006). What we have learned: RC28's contributions to knowledge about social stratification. *Research in Social Stratification and Mobility*, 24(1):1–20.

Huang, M.-H. (2020). Compensatory advantage and the use of out-of-school-time tutorials: A cross-national study. *Research in Social Stratification and Mobility*, *66*, 100472. https://doi.org/https://doi.org/10.1016/j.rssm.2020.100472

Jackson, M. (2013). How is inequality of educational opportunity generated? The case of primary and secondary effects, in: M. Jackson (ed.): *Determined to succeed? Performance versus choice in educational attainment*, Stanford: Stanford University Press, Chapter 1 (pp. 1-33).

Jackson, M. and Jonsson, J.O. (2013). Why Does Inequality of Educational Opportunity Vary Across Countries? Primary and Secondary Effects in Comparative Context, in: M. Jackson (ed.): Determined to succeed? Performance versus choice in educational attainment, Stanford: Stanford University Press, Chapter 11 (pp. 306-337).

Jæger, M. M. (2011). Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data. *Sociology of Education*, 84(4), 281–298. doi:10.1177/0038040711417011

Jæger, M. M., and Breen, R. (2016). A Dynamic Model of Cultural Reproduction. American *Journal of Sociology*, 121(4), 1079–1115.

Jæger, M. M., and Møllegaard, S. (2017). Cultural capital, teacher bias, and educational success: New evidence from monozygotic twins. *Social Science Research*, 65, 130–144. doi:https://doi.org/10.1016/j.ssresearch.2017.04.003

Jencks, C., and Laura Tach (2006). Would Equal Opportunity Mean More Mobility? In Stephen Morgan, David Grusky, and Gary Fields (editors). *Mobility and Inequality: Frontiers of Research from Sociology and Economics.* Stanford: Stanford University Press.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review.* 39, 1–123. doi:10.17763/haer.39.1.l3u15956627424k7

Johnson, W., Turkheimer, E., Gottesman, I. I., and Bouchard, T. J. (2009). Beyond heritability: Twin studies in behavioral research. *Current Directions in Psychological Science*, 18(4), 217–220.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

Kaiser, T. (2016). Socioeconomic status, parenting and conscientiousness: the unequal development of children's focus, *Journal of Family Studies*, DOI: 10.1080/13229400.2016.1152195

Kaiser, T., and Diewald, M. (2014a). Ordentliche Arbeiterkinder, konzentrierte Mittelschicht-kinder? Die ungleiche Entwicklung von Gewissenhaftigkeit im frühen Kindesalter [Tidy working-class children, focused middle-class children? The unequal development of conscientiousness in early childhood]. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66(2), 243–265.

Kendler, K. S., et al. (2015). Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proceedings of the National Academy of Sciences*, 112 (15), 4612–4617. doi:10.1073/pnas.1417106112

Knol, M. J., and VanderWeele, T. J. (2012). Recommendations for presenting analyses of effect modification and interaction. *International Journal of Epidemiology*, *41*(2), 514–520. https://doi.org/10.1093/ije/dyr218

Knopik, V. S.; Neiderhiser, J. M; DeFries, J. C, and Plomin, R. (2017). *Behavioral genetics* (7th ed.). New York, NY: Worth.

Knudsen, E. I.; J. J. Heckman, J. L. Cameron, and J. P. Shonkoff (2006): Economic, neurobiological, and behavioral perspectives on building America´s future workforce, in: *Proceedings of the National Academy of Sciences* (PNAS), July 5, 2006, No. 27: 10155-10162.

Kohler, H. P., Behrman, J. R. and Schnittker, J. (2011). Social science methods for twins data: Integrating causality, endowments, and heritability. *Biodemography and Social Biology*, 57(1), 88–141.

Korber M and Oesch D (2019). Vocational versus general education: Employment and earnings over the life course in Switzerland. *Advances in Life Course Research*, 40: 1–13.

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J.-B., et al. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences*, 111(42), 15273–15278. doi: 10.1073/pnas.1408777111

Kulic, N., Skopek, J., Triventi, M., and Blossfeld, H.-P. (2019). Social Background and Children's Cognitive Skills: The Role of Early Childhood Education and Care in a Cross-National Perspective. *Annual Review of Sociology*, *45*(1), 557–579.

Landes, X., and Nielsen, M. (2012). Intra-Family Inequality and Justice. *Dialogue*, 51(3), 437-466. doi:10.1017/S0012217312000698

Lareau, A. (2003). *Unequal Childhoods: Class, Race, and Family Life*. Berkeley: University of California Press.

Lareau, A. (2015). Cultural Knowledge and Social Inequality. *American Sociological Review*, 80(1): 1–27.

Layte, R. (2017). Why Do Working-Class Kids Do Worse in School? An Empirical Test of Two Theories of Educational Disadvantage, *European Sociological Review*, 33(4):489-503, https://doi.org/10.1093/esr/jcx054

Liu, A. (2019). Can non-cognitive skills compensate for background disadvantage ? – the moderation of non-cognitive skills on family socioeconomic status and achievement during early childhood and early adolescence. *Social Science Research*, *83*(April), 102306.

Loehlin, J. C., Harden, K. P., Turkheimer, E. (2009). The effect of assumptions about parental assortative mating and genotype–income correlation on estimates of genotype–environment interaction in the national merit twin study. *Behavior Genetics*, 39(2), 165–169.

Lucas, S. R. (2017). An Archaeology of Effectively Maintained Inequality Theory. *American Behavioral Scientist*, 61(1): 8–29.

Lugo-Gil, J., and Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Development*, 79(4), 1065-1085. DOI: 10.1111/j.1467-8624.2008.01176.x

Marshall, G., Swift, A., and Roberts, S. (1997). *Against the Odds? Social Class and Social Justice in Industrial Societies*. Oxford: Oxford University Press.

Mcewen, C. A., and Mcewen, B. S. (2017). Social Structure, Adversity, Toxic Stress, and Intergenerational Poverty: An Early Childhood Model, *Annual Review of Sociology*, (April), 1–28.

McGue, M., Osler, M., and Christensen, K. (2010). Causal Inference and Observational Research: The Utility of Twins. *Perspectives on Psychological Science*, 5(5): 546–556

Meghir, C. and Palme, M. (2005). Educational Reform, Ability, and Family Background. *American Economic Review*, 95(1):414-424.

Mijs, J.J.B (2019). The paradox of inequality: income inequality and belief in meritocracy go hand in hand, *Socio-Economic Review*, https://doi.org/10.1093/ser/mwy051

Mijs, J.J.B. (2016). The Unfulfillable Promise of Meritocracy: Three Lessons and Their Implications for Justice in Education. *Social Justice Research*, **29,** 14–34. https://doi.org/10.1007/s11211-014-0228-0

Need, A. and de Jong, U. (2001). Educational differentials in the Netherlands: testing rational action theory. *Rationality and Society*, 13: 71–98.

Neisser, Ulric, Gwyneth Boodoo, Thomas J. Bouchard, Jr., A. Wade Boykin, Nathan Brody, Stephen J. Ceci, Diane F. Halpern, John C. Loehlin, Robert Perloff, Robert J. Sternberg and Susana Urbina. (1996). "Intelligence: Knowns and Unknowns." *American Psychologist* 51:77-101.

Nelson, C. A., and M.A. Sheridan. (2011). Lessons from Neuroscience Research for Understanding Causal Links between Family and Neighborhood Characteristics and Educational Outcomes. *Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances*, edited by G. J. Duncan and R. J. Murnane, 27–46. New York: Russell Sage.

Nicoletti, C., and Tonei, V. (2020). Do parental time investments react to changes in child's skills and health? *European Economic Review*, *127*, 103491. https://doi.org/https://doi.org/10.1016/j.euroecorev.2020.103491

Nielsen, F. (2006). Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling. *Social Forces* 85(1):193–216.

Nielsen, F. (2016). The Status-Achievement Process: Insights from Genetics. *Frontiers in Sociology*. Retrieved from https://www.frontiersin.org/article/10.3389/fsoc.2016.00009

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., and Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *The American Psychologist*, 67(2), 130–59. doi:10.1037/a0026699

OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/9789264073234-en.

Okbay, A., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, 533(7604), 539–542.

Parsons, Talcott (1951). *The Social System.* Illinois: Glencoe.

Passaretta, Giampiero, Jan Skopek, and Thomas van Huizen (2020). "To What Extent Is Social Inequality in School-age Achievement Determined Before and During Schooling? A Longitudinal Analysis in Three European Countries." *SocArXiv.* April 3. doi:10.31235/osf.io/yqt6n

Pearl, J., and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*, Hachette UK.

Pekkala Kerr, S., Pekkarinen, T., and Uusitalo, R. (2013). School Tracking and Development of Cognitive Skills. *Journal of Labor Economics*, 31(3): 577–602.

Pekkarinen, T., Uusitalo, R., and Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, 93(7-8): 965–973.

Pfeffer, F.T. (2008). Persistent inequality in educational attainment and its institutional context. *European Sociological Review*, 24: 543–565.

Piketty, T. (2020). *Capital and Ideology.* Cambridge: Belknap Press

Piketty, T. (2000). Theories of persistent inequality and intergenerational mobility. In A. B. Atkinson, and F. Bourguignon (Eds.), *Handbook of income distribution* (pp. 429–476). Amsterdam: Elsevier

Protzko, J. (2015). The environment in raising early intelligence : A meta-analysis of the fadeout effect. *Intelligence*, 53(November), 202–210. doi:10.1016/j.intell.2015.10.006

Raftery, A.E. and Hout, M. (1993). Maximally Maintained Inequality. Expansion, Reform, and Opportunity in Irish Education 1921–75. *Sociology of Education*, 66: 41–62.

Rauscher, E. (2016). Does Educational Equality Increase Mobility? Exploiting 19th Century U.S. Compulsory Schooling Laws. *American Journal of Sociology*, 121(6): 1697-1761.

Reardon, S. (2011). "The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations." Pp. 91–116 in *Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances*, edited by Murnane, R., Duncan, G. New York: Russell Sage Foundation.

Rietveld, C. A., S. E. Medland, J. Derringer, and J. Yang (2013). "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science* 340(6139):1467–71. http://dx.doi.org/10.1126/science.1235488.

Rindermann, H. (2007). The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS and IQ-Tests across Nations. European *Journal of Personality* 21(5):667–706.

Ritchie, S. J., and Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? A Meta-Analysis. *Psychological Science*, *29*(8), 1358–1369.

Roemer, J. (1998). Igualdad de Oportunidades. *Isegoría*, nº 18, pp. 71-87.

Roemer, J. (2012). What Is the Justification of Studying Intergenerational Mobility of Socioeconomic Status? In Ermisch J., Jäntti M., and Smeeding T. (Eds.), *From Parents to Children: The Intergenerational Transmission of Advantage* (pp. 482-488). Russell Sage Foundation.

Rowe, David C., Wendy J. Vesterdal and Joseph L. Rodgers. (1999). Herrnstein's Syllogism: Genetic and Shared Environmental Influences on IQ, Education, and Income. *Intelligence* 26:405-23.

Rudolphi, F. (2013). Ever declining inequalities? Primary and secondary effects in Sweden in the transition to upper secondary and higher education for students born in 1972-1990. In: Jackson, M. (Ed.), *Determined to Succeed? Performance, Choice and Education.* Stanford: Stanford University Press.

Salazar, L., Cebolla-Boado, H., and Radl, J. (2019). Educational expectations in the great recession: has the impact of family background become stronger? *Socio-Economic Review.* https://doi.org/10.1093/ser/mwy046

Sapolsky, R. M. (2017). *Behave: The Biology of Humans At Our Best and Worst.* New York, New York: Penguin Press.

Saunders, P. (1996). Social Mobility in Britain: an Empirical Evaluation of Two Competing Explanations, *Sociology*, 31 (2), pp. 261-288.

Schindler, H.S., et al. (2015). Maximizing the potential of early childhood education to prevent externalizing behavior problems : A meta-analysis. *Journal of School Psychology*, *53*(3):243–263.

Schneider, D., Hastings, O., and LaBriola, J.. (2018). Income Inequality and Class Divides in Parental Investments. *American Journal of Sociology*, 83 (3), pp. 475–507.

Shanahan, M., et al. (2014). Personality and the reproduction of social class. *Social Forces*, *93*(1), 209–240.

Shavit, Y., and Blossfeld, H.P. (1993). *Persisting inequality: changing educational attainment in thirteen countries*, Oxford: Westview Press.

Shavit, Y and Müller W (eds) (1998). *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations.* Oxford: Oxford University Press.

Sieben, I., and de Graaf, P. (2003). The total impact of the family on educational attainment. (2003). European Societies, 5(1), 33–68.

Skopek, J.; N. Kulic, M. Triventi, and H.-P. Blossfeld (2016). Childcare, early education and compensation of educational disadvantage – Evidence from a multidisciplinary and international project, In: H.-P. Blossfeld, N. Kulic, J. Skopek, and M. Triventi (eds.): *Childcare, early education and social inequality. An international perspective*, Vol. 4 eduLIFE Lifelong Learning Series, Cheltenham (UK) and Northampton (MA, USA): Edward Elgar (1-27).

Sniekers, S. et al., (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature Genetics*, 49(7), 1107–1112. http://dx.doi.org/10.1038/ng.3869

Spence, M. (1973). Job Market Signalling. *The Quarterly Journal of Economics*, 87(3): 355 374.

Spinath, B. (2005). Development and modification of motivation and self-regulation in school contexts: Introduction to the special issue. *Learning and Instruction, 15*(2), 85–86.

Spinath, B., and Spinath, F. M. (2005). Development of self-perceived ability in elementary school: the role of parents' perceptions, teacher evaluations, and intelligence. *Cognitive Development, 20*(2), 190–204.

Südkamp, A., Kaiser, J., and Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*(3), 743-762.

Sullivan, A. (2002). Bourdieu and Education: How Useful is Bourdieu's Theory for Researchers? *Netherlands Journal of Social Sciences*, 38(2): 144-166.

Swift, A. (2003). *How not to be a Hypocrite: School Choice and the Morally Perplexed Parent.* New York: Routledge.

Swift, A. (2004). Would Perfect Mobility be Perfect? *European Sociological Review*, 20(1):1-11.

Swift, A. (2005) Justice, luck, and the family: the intergenerational transmission of economic advantage from a normative perspective. In: Bowles, Samuel and Gintis, Herbert and Osbourne Groves, Melissa, (eds.) *Unequal chances: family background and economic success.* New York: Princeton, N.J. ; Oxford : Russell Sage Foundation: Princeton University Press, pp. 256-276.

Torche, F. (2015) Intergenerational Mobility and Equality of Opportunity. *European Journal of Sociology*, 56(3): 343–371.

Treiman, D.J. (1970). Industrialization and Social Stratification. In E. O. Laumann (Ed.), *Social Stratification: Research and Theory for the 1970s*, Indianapolis: Bobbs Merrill.

Triventi, M. (2019). Are Children of Immigrants Graded Less Generously by their Teachers than Natives, and Why? Evidence from Student Population Data in Italy. *International Migration Review.*

Triventi, M., J. Skopek, N. Kulic, S. Buchholz, and H.-P. Blossfeld (2016): Varieties of secondary education models and social inequality – Conclusions from a large-scale international comparison. In: In: H.-P. Blossfeld, S. Buchholz, J. Skopek and M. Triventi (eds.): *Models of secondary education and social inequality. An international comparison*, Vol. 3 eduLIFE Lifelong Learning Series, Cheltenham (UK) and Northampton (MA, USA): Edward Elgar (377-400).

Tropf, F. C., et al. (2017). Hidden heritability due to heterogeneity across seven populations. *Nature Human Behaviour*, 1: 757–765. doi:10.1038/s41562-017-0195-1

Tucker-Drob, E. M., and Harden, K. P. (2012a). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene-environment transactions. *Developmental Science*, 15, 250-259. doi: 10.1111/j.1467-7687.2011.01121.x

Tucker-Drob, E. M., Briley, D. A., and Harden, K. P. (2013). Genetic and environmental influences on cognition across development and context. *Current Directions in Psychological Science*, 22, 349-355. doi: 10.1177/0963721413485087

Turkheimer, E. (2011). Still Missing. *Research in Human Development*, 8(3-4), 227–241.

Turkheimer, E., and Harden, K. P. (2014). Behavior genetic research methods: Testing quasi-causal hypotheses using multivariate twin data. In H. T. Reis and C. M. Judd (Eds.), *Handbook of research methods in personality and social psychology* (2nd ed., pp. 159-186). Cambridge, England: Cambridge University Press.

Turkheimer, Eric. (2000). Three Laws of Behavior Genetics and What They Mean. *Current Directions in Psychological Science* 5:160-64.

Van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G., and Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. Nature Reviews *Genetics*, 13(9), 640–653. http://dx.doi.org/10.1038/nrg3243

Waller, J. (2012). Commentary: The birth of the twin study—a commentary on Francis Galton's 'The History of Twins'. *International Journal of Epidemiology*, 41:917–919.

Watts, T. W., Duncan, G. J., and Quan, H. (2018). Revisiting the Marshmallow Test: A conceptual replication investigating links between early delay of gratification and later outcomes. *Psychological Science*, 29(7), 1159-1177.

Wheen, F. (2001). *Satirical fiction is becoming Blair's reality.* The Guardian, February 14. Retrieved March 23th, 2017. (http://www.theguardian.com/theguardian/2001/feb/14/features11.g21).

Young, M. (1958). *The rise of the meritocracy, 1870-2033. An Essay on Education and Equality.* Harmondsworth: Penguin.

Chapter II
Birth Weight and Skill Formation: Biological Destiny or Parental Response?

Carlos J. Gil-Hernández

Marco Cozzani

Fabrizio Bernardi

Abstract

Birth weight (BW) is a key predictor of child development and socioeconomic attainment later in life. However, the consequences of BW for status-attainment are not biological destiny. Educational interventions targeted at disadvantaged families are successful in offsetting the negative consequences of prenatal health shocks. Similarly, parents respond with investments to children's birth endowments, influencing their later skill formation. This article tests whether high-socioeconomic status (SES) families are able to neutralise/compensate for prenatal health shocks thanks to their large pool of economic and cultural resources. A socioeconomic gradient in the effect of BW on skill formation may contribute to the persistence of early SES-gaps over the life-course. We study two cohorts at ages 5 and 11, drawing from the German *Twin Life Study*. We implement twin fixed-effects models to estimate the causal effect of BW by exploiting random sources of variation in intrauterine growth between twins. Results show that lower-BW co-twins have worse academic performance and more behavioural problems than their heavier-BW co-twins. At age 5, we observe a causal effect of BW on academic performance and behavioural problems that fades away for high-SES children at age 11. We argue that this compensatory pattern at age 11 may be explained by high absolute levels of investments by high-SES families (e.g., cultural activities and warmth), but not by their relative allocation of investments within families. Thus, we argue that biology is not destiny because (enriched) social environments may offset the detrimental effect of BW on skill formation.

1. Introduction

Health at birth is a crucial circumstance for skill formation during sensitive stages of child development (Heckman 2007). At the starting gate of life, low birth weight (LBW; below 2,500 grams) represents a serious health risk affecting around 6–16% of single births and 50% of multiple births of the population of newborns worldwide (Blencowe et al. 2019). Birth weight (BW) is widely studied as a proxy for children's *in utero* environment, perinatal health and developmental potential (Torche and Conley 2016). Birth weight is also a good predictor of educational and socioeconomic attainment later in life (Black, Devereux and Salvanes 2007).

However, the consequences of BW for socioeconomic attainment are not biological destiny (Conley and Bennett 2000). Educational interventions targeted at LBW children in disadvantaged families were successful in offsetting the negative impact of prenatal health shocks on skill formation (McCormick et al. 2006; Walker et al. 2010). There is also ample evidence that parental involvement is related to children's well-being and educational success (Lugo-Gil and Tamis-LeMonda 2008). Thus, parents may offset the negative consequences of prenatal health shocks (Attanasio et al. 2020). However, the level and quality of postnatal parental investments, and their capacity to counterbalance (or strengthen) health shocks, depends on the economic and cultural resources of the families (Torche 2018).

In this article, we study the effect of BW on educational outcomes in Germany at two points in children's development (ages 5 and 11), how it differs by family socioeconomic status (SES), and the mechanisms underlying its heterogeneity by family SES. We address the following research questions: (1) Is BW associated with children's cognitive and non-cognitive skills at 5-11 years old? (2) Does the effect of BW on academic skills vary by parental SES? (3) If so, are high-SES families more likely to reduce (or neutralise) the effect of BW on academic skills than low-SES families?

Previous research has widely investigated the relationship between BW and children's socioeconomic outcomes (Almond, Currie and Duque 2018) but the number of studies that have explicitly focused on the heterogeneity of these relationships by parental SES is still limited, and results are mixed (Torche and Conley 2016). Thus, in answering the aforementioned research questions, we provide two contributions to an emerging interdisciplinary literature at the crossroads between social stratification; developmental psychology, epidemiology, and human capital formation.

Firstly, we add to the literature because we not only study SES heterogeneity in the effect of BW but further explore two (alternative) mechanisms explaining this potential stratification: (1) parental response or allocation of investments among children within the same family (mediation); and (2) the absolute level of resources between families (moderation).

On the one hand, we assess whether parental allocation of investments within families mediates the effect of BW on later academic outcomes (Lynch and Gibbs 2016). Families tend to *respond* to their children's early observable endowments by investing more (reinforcement) or less (compensation) in the more-endowed sibling (Behrman et al. 1982). In particular, we test the hypothesis that high-SES families compensate while low-SES families reinforce for BW differences in their allocation of investments due to resource constraints (Conley 2008). Current literature offers mixed results on the parental response (Almond, Currie and Duque 2018).

On the other hand, we look at the role of the absolute level of parental investments between families in moderating the effect of BW on academic outcomes. The large absolute level of economic and cultural resources at the disposal of high-SES families allows them to deploy high-quality educational investments through parenting and schooling (Bernardi 2014; Torche 2018). Thus, if parental resources and investments are above a critical threshold for child development, the effect of BW may vanish independently of how parents allocate investments within families. We discuss the general implications of these arguments for research design based on siblings and twins models, which are common practice in this area of research (Almond and Mazumder 2013).

Secondly, we contribute to the literature on the multidimensionality of skill formation by assessing the effect of BW on three key outcomes for learning and educational success (Smithers et al. 2018): behavioural problems; cognitive ability; and academic performance in mathematics and language. These skills are plausible mechanisms accounting for the long-term effect of BW on educational attainment and labour market outcomes found in previous literature (Black et al. 2007).

To test our research questions, we use data from the first wave of the *Twin Life – Genetic and Social Causes of Life Chances Study* carried out in 2014/2015, a register-based representative survey of the German population that comprises two cohorts (n=4,096) of same-sex monozygotic (MZ) and dizygotic (DZ) twins at ages 5 (born 2009/10) and 11 (born 2003/04).

We implement twin fixed-effects models to estimate the causal effect of BW by exploiting random sources of variation in intrauterine growth between twins (e.g., twin-specific placental position) raised in the same family. In this way, we can rule out potential confounders that affect both BW and later educational outcomes (e.g., parental SES, risky behaviours, body mass index, and mother's age) (Currie 2011). Birth weight differences between twins can then be conceived as the outcome of a natural experiment (Torche and Conley 2016). However, it should be noted that twin fixed-effects rule out the main cause of LBW among singletons in Western societies, prematurity, with about 70% of LBW newborns in Western countries being pre-term but normal for gestational age. In the remainder of this paper, we will discuss the trade-off between causal identification and external validity given the particular aetiology of BW among twins and their left-skewed distribution, with a high prevalence of LBW and prematurity at about 50%.

The German case is particularly relevant for testing theories of skill-formation and intergenerational educational inequality given its early ability-tracking educational system, which funnels pupils into academic or vocational training pathways at age 10. Thus, a socioeconomic gradient in the effect of BW on skill formation may contribute to the persistence of early SES-gaps over the life-course (Skopek and Passaretta 2018).

2. Theoretical Background

In this section, we firstly build on the human capital formation literature to present a general theoretical framework on the relationships between birth endowments, parental investments, and children's academic skills. Secondly, we explain how these associations are stratified by parental SES. Finally, we explore two alternative mechanisms accounting for the heterogeneous effect of BW on academic skills by parental SES: (1) within-family allocation of investments; and (2) between-family differences in absolute investments.

2.1. Human Capital Formation: Birth Endowments, Investments and Skills

Early childhood is a sensitive period for child development, and health at birth may be a fundamental determinant. This early stage can be understood as a window of developmental opportunity (e.g., central nervous system; brain malleability) in which health shocks may have strong and long-term consequences for future skill formation (Knudsen et al. 2006).

According to theories of early human capital formation (Heckman 2007), the current stock of children's human capital is a dynamic function of genetic endowments, previous skills, and parental investments. Skills are multidimensional and embrace health, cognitive, and non-cognitive (e.g., socio-emotional skills) attributes (Francesconi and Heckman 2016). Parental investments are also multifaceted (e.g., money, time, knowledge), depend on resource constraints and preferences (e.g., altruism; inequity aversion), and mainly encompass parental involvement and schooling quality.

As illustrated in Figure 1, the level of skills at a certain stage of child development (BW as birth endowments) has a direct effect on the level of skills at a subsequent stage (arrow a), along with an indirect effect through parental environment (arrows b*c): parental investments may mediate (and also moderate; arrow d) the association between children's birth endowments and later skill formation. In the next subsections we describe in detail each association displayed in Figure 1.

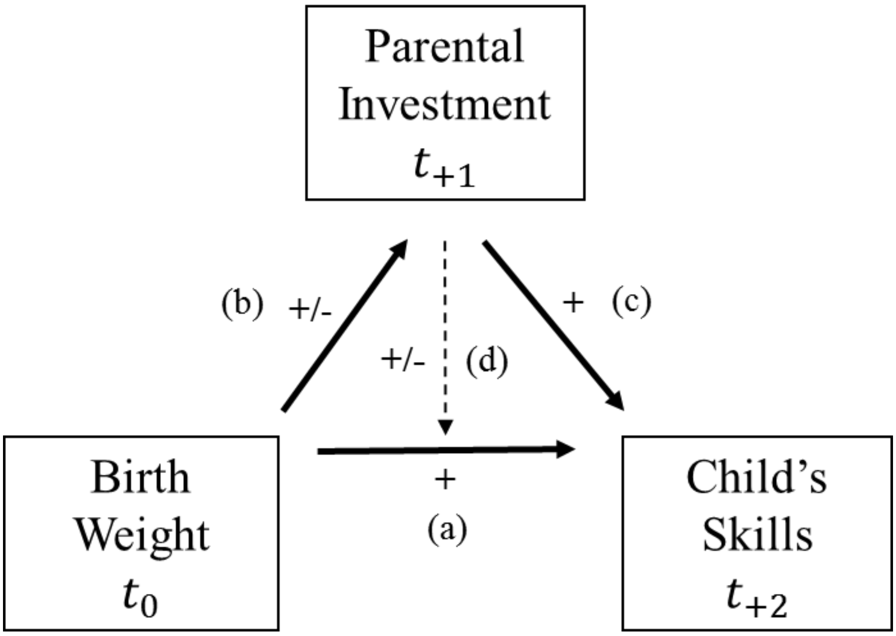


Figure 1. Associations between: (a) BW at t_0 and academic skills at t_{+2}; (b) BW at t_0 and parental investment at t_{+1}; (c) parental investment at t_{+1} and academic skills at t_{+2}; (d) BW at t_0 and academic skills at t_{+2} over parental investment at t_{+1} (moderator).

2.1.1. Birth Weight and Skill Formation

Birth weight is widely considered as a proxy for perinatal health and developmental potential in early childhood[23] (Conti et al. 2018). Yet, what does BW really capture about the prenatal environment? BW mainly reflects the level of nutrition (both lean mass and fat mass)[24] *in utero* during the last weeks of gestation. Patterns of foetal growth in the third or last term of gestation are related to mental and physical health conditions such as brain development (volume), asthma and hyperactivity (Conti et al. 2018). Thus, BW can be considered as a proxy for the late prenatal environment, while other direct measures of *in utero* foetal size, such as head circumference, abdominal circumference and femur length, better capture the whole foetal trajectory or the body composition (fat). However, these latter measures are not commonly available in observational data.

Accumulating research shows the negative long-term impact of LBW and very low birth weight (VLBW; < 1,500 grams) on cognitive and non-cognitive skills. In a meta-analysis of the effect of being born very preterm (≤ 33 weeks of gestation) and VLBW, Aarnoudse-moens et al. (2009) found moderate-to-severe deficits in academic performance in mathematics and reading, and behavioural problems (e.g., attention and internalising problems). Most research on the effects of BW on early neuropsychological development is focused on preterm (< 37 weeks of gestation), very preterm and VLBW infants (Sripada et al. 2018). However, variation in the range of a normal BW (NBW; ≥ 2,500 grams) is, to a lesser extent, also predictive of cognitive and behavioural outcomes (Walhovd et al. 2012). Sripada et al. (2018) argue that "the preterm behavioural phenotype has been described as anxious and inattentive, rather than hyperactive or disruptive, which may also mean that their cognitive difficulties may not be as readily visible in a classroom setting."

Among LBW thresholds, magnetic resonance imaging has shown impairments in the growth of certain key brain structures (i.e., caudate nuclei, pertaining to learning and memory, and the hippocampus) associated with learning difficulty, attention deficit, and dyspraxia (developmental coordination disorder) (Abernethy et al. 2002). Moreover, VLBW is related to disruptions of cortical and subcortical circuits connecting the frontal, striatal and thalamic regions of the brain (Aarnoudse-Moens et al. 2009). These disruptions have negative consequences for the executive functions of the brain, such as inhibitory control, working

[23] Other measures of birth endowments are birth length, head circumference and APGAR scores.
[24] Given that BW also captures fat mass, which increases the likelihood of macrosomia and early delivery, it is important to adjust BW for gestational age.

memory, cognitive flexibility, and planning, that are considered as foundational for academic performance.

2.1.2. Parental Investments and Skill Formation

A stream of research building on human capital formation models assesses whether parental investments can offset the negative impact of prenatal health shocks on skill formation later in life. As put by Aizer and Currie (2014, 859): "Poor health at birth will only have lasting effects if parents are unable or unwilling to offset its impacts through postnatal parental investments." Likewise, educational interventions targeted at LBW children from disadvantaged backgrounds in the USA and Jamaica have been successful in offsetting the negative impact of prenatal health shocks in later skills' development through intensive psychosocial stimulation (McCormick et al. 2006; Walker et al. 2010).

There are two main features of positive parenting stressed by developmental psychologists that enhance children's development of language, cognition, socio-emotional skills and school readiness (arrow c in Figure 1): cognitive stimulation (e.g., learning-promoting activities); and positive parents-child interactions, as usually measured by sensitivity (e.g., emotional support) and warmth (e.g., affection and respect) (Attanasio et al. 2015; Lugo-Gil and Tamis-LeMonda 2008, 1066).

Therefore, parental investments or involvement may intervene in the association between BW and later skills by mediation (arrows b*c in Figure 1) or moderation (arrow d in Figure 1). First, parents may differentiate their investments depending on the endowments of their children (arrow b in Figure 1) (Grätz and Torche 2016). For instance, if NBW children receive more and better parental investments than LBW children (positive arrow b in Figure 1), this situation would reinforce the effect of BW on later skill differences (Lynch and Gibbs 2016). If, on the contrary, LBW children receive more and better parental investments than NBW children (negative arrow b in Figure 1), this allocation of investments may neutralise or compensate for the impact of BW on later skill disparities.

Second, high levels of parental investments may be especially effective and pay off for children with more initial endowments (positive arrow d in Figure 1) (Aizer and Cunha 2012). Normal birth weight children tend to develop earlier and would then benefit more from the stimulation by their parents than LBW children, so reinforcing later differences in academic skills between them. Alternatively, as LBW children usually have a greater prevalence of impairments and more room for development in comparison to NBW children, the former would benefit to a greater extent from high levels of parental investments (negative arrow d in Figure 1), so that skill differences may eventually disappear between both groups of children.

2.2. The Stratification of Health Shocks and Investments by Parental SES

The level and quality of parental investments/involvement and, consequently, their capacity to reduce or neutralise health shocks, depends on the economic and cultural resources of the families (Torche 2018). As Aizer and Currie (2014, 860) argue: "Children with poorer initial health endowments typically receive fewer postnatal investments, and the investments that they do receive may be less effective. This mechanism can potentially explain the considerable persistence of *in utero* conditions on later offspring outcomes. It can also explain why the long-term impact of low birth weight is greater when children are born into poverty."

Sociological theories on accumulative (dis)advantage also highlight that high-SES families have more resources and incentives (e.g., intergenerational status reproduction) than low-SES families to compensate for negative shocks. In particular, the *compensatory advantage hypothesis* argues that the life-course trajectories of kids/pupils from privileged backgrounds are less dependent on prior negative outcomes or disadvantageous traits due to proactive parental involvement (Bernardi 2014).

Indeed, the quality and quantity of parental investments vary across the socioeconomic strata (Aizer and Cunha 2012). As we pointed out above, there are two main features of positive parenting enhancing children's well-being and academic skills that also vary by SES: cognitive stimulation, and positive parent-child interactions (Lugo-Gil and Tamis-LeMonda 2008).

With respect to the stimulation component, sociological research drawing from Bourdieu's cultural capital[25] theory and ethnographic research (Lareau 2003) looked at parental time use with children in cultural activities (e.g., reading, playing), its impact on educational performance and stratification by SES (Jæger 2011). According to Lareau (2003, 238), middle-upper class families follow a strategy of *concerted cultivation* in which parents actively assess

[25] Cultural capital is usually conceptualised in three dimensions: Institutionalised cultural capital (formal): educational credentials; embodied cultural capital (socialisation): *habitus*; and objectivised cultural capital (material cultural resources): books, pieces of art, etc.

their children's skills and opinions, reason with them, and make sustained efforts to stimulate and cultivate their cognitive and social skills.

By the same token, building on the classic human capital theory (Becker and Tomes 1976), economists tend to emphasise the level of parental economic resources (i.e., earnings) and the different investments of money and time that they allow. Wealthy families can afford sustained financial and time investments to enhance the learning opportunities of their children (i.e., enriching educational toys, private tutors and schools, summer camps, extracurricular activities, bedtime stories).

Recent developments of the human capital theory (Francesconi and Heckman 2016) emphasise the importance of separating parental resources into financial investments and parental education. As Francesconi and Heckman point out (2016:11): "As it is imprecise to proxy human capital by scores on intelligence quotient (IQ) or achievement tests, it is inadequate to measure parental investment only in terms of financial expenditures on the child [...] Levels of permanent income are highly correlated with family background factors such as parental education and maternal ability, which, when statistically controlled for, largely eliminate the gaps across income classes."

Regarding positive parent-child interactions, extensive research from developmental psychopathology and neurobiology show how, in the first years of life, parental warmth functions as a protective factor for behavioural and cognitive problems in children (Laucht, Esser and Schmidt 2001; Tully et al. 2004). Its theoretical foundation lies in the "attachment theory", which posits that children need a strong bond with the main caregiver (i.e., maternal responsiveness and warmth) to build a *secure base* upon which to explore the world to eventually engage in adult-supervised learning experiences. The evidence supports the association between high-quality parents-child interactions and children's brain development, cognitive ability and educational performance (Ranson and Urichuk 2008).

More interestingly for the purpose of this article, Tully et al. (2004:2) studied a group of LBW children and found that high levels of maternal warmth moderate the risk of long-term cognitive and behavioural problems. Likewise, Laucht, Esser and Schmidt (2001) carried out a longitudinal study (across ages 2-8) looking at the effects of mothers' emotional responsivity on attention problems among LBW children. They found a moderating effect of maternal responsivity by which LBW was less detrimental for attention problems among those children whose mothers were emotionally responsive.

Positive parents-child interactions also tend to vary by SES (Farah et al. 2008). For instance, disadvantaged families who cope with poverty and deprivation (McEwen and McEwen 2017) on a daily basis tend to work extended hours, have unstable schedules, and long commuting times, so increasing their exposure to stress, marital conflict, and related psychological problems. These constraints lead to fewer resources and less time to invest in their children, negatively affecting parenting quality (e.g., warmth, monitoring, and stimulation) (Layte 2017). This is especially the case in liberal welfare states such as the US, while in the European context of more comprehensive social policies, the effects of poverty on children's development are more moderate. This suggests that income is more important when it leads to actual poverty.

Given the aforementioned differences by SES in resources and parenting quality, we expect an observed pattern of *compensatory advantage* by which the (negative) effect of BW on academic skills is reduced (or neutralised) for high-SES families (a $\approx$ 0 in Figure 1) with respect to low-SES families (positive arrow a in Figure 1). In the next subsections, we elaborate on the heterogeneous effect of BW across the socioeconomic gradient by exploring two mechanisms: (1) parental response or allocation of investments among children within the same family (mediation); and (2) the absolute level of resources between families (moderation).

2.2.1. Within-Family Allocation of Investments

According to classic microeconomics' theories of intra-household resource allocation (Becker and Tomes 1976), parents allocate resources among siblings depending on three factors: (1) preferences (e.g., altruism; inequity aversion); (2) perception of the children's endowments (e.g., information on children's skills and returns on investment); and (3) budget constraints. Intra-family models assess whether parents allocate investments among their children in a neutral, reinforcing or compensatory way.

Parents may be driven by efficiency concerns, maximising returns and investing more resources in the more-endowed child (*reinforcing parental response*) (Becker and Tomes 1976). On the contrary, if parents are averse to within-family inequity, they will invest more in the less-endowed child (*compensating parental response*) (Behrman et al. 1982). In a nutshell, for this mediation mechanism to hold, we must find an initial link between BW and parental investment (b ≠ 0 in Figure 1) (Lynch and Gibbs 2016).

Building upon these microeconomic models, sociological theories emphasise different parental strategies of resource allocation across the socioeconomic hierarchy (Conley 2008). The absolute level of family resources constrains parental *preferences* for intra-family resource allocation: if low-SES do not have enough resources to compensate for differences in endowments among their children, they will "bet" on the more-endowed child to get the highest return to their investments (positive arrow b in Figure 1). In turn, as high-SES parents are freer from resource constraints, they will compensate for differences in early endowments among their kids by investing more in the less-endowed child (negative arrow b in Figure 1).

2.2.2. Absolute Investments Between Families

High-SES families are more likely to neutralise the negative consequences of health shocks than low-SES families thanks to their larger pool of economic and cultural resources (Bernardi 2014; Conley 2008; Torche 2018). This absolute pool of resources would allow high-SES families to deploy higher-quality and more-effective investments than low-SES families through parenting and schooling. Thus, relative differences in parental investments among children in the same family may be trivial if the absolute level of resources lies above a critical threshold for child development, and vice versa.

Theories on human capital formation describe the functional form that relates parental investments with children's skills as concave (Francesconi and Heckman 2016): a steeper positive function at low levels of investments that, over a certain threshold, levels out. This concave functional form implies that *a one unit-increase* in parental investments has a stronger impact on children's skills at low absolute levels of investments. In this scenario, parental allocation of investments that reinforce siblings' differences in early endowments is only salient at low levels of absolute investments (resources), whereas they become irrelevant over a given threshold.

For all of the above, the marginal effect of BW on later skills is expected to decrease at high absolute levels of parental investments (negative arrow d in Figure 1), independently of how parents allocate resources within-families. Lighter-BW siblings may benefit more from high levels of absolute investments than their heavier-BW siblings, hence reducing or neutralising their relative differences in later skills, as the former have more room for skill development, playing catch up.

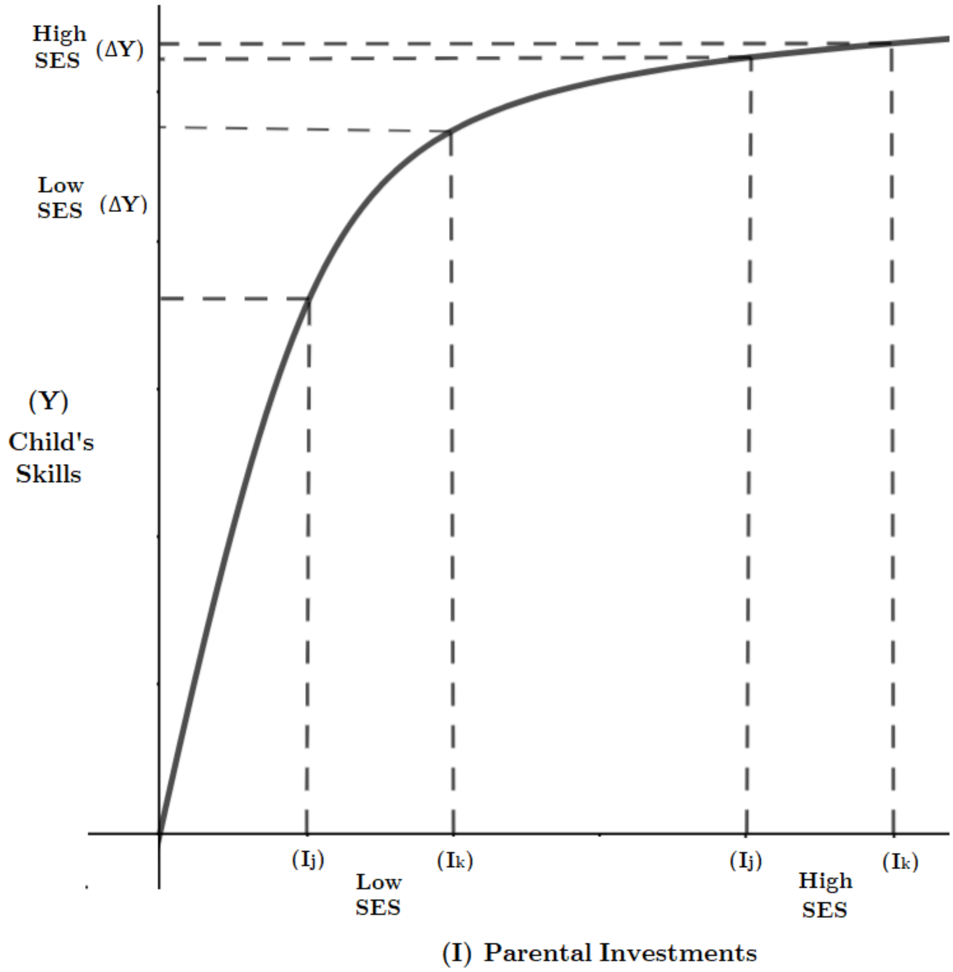


Figure 2. Theoretical functional form of the relationship between parental investments and children's academic skills by parental SES

Notes: I_j= investment in sibling j in family i; I_k= investment in sibling k in family i;
ΔY Skills' differences between sibling j and sibling k in family i.

3. Previous Research

Since the early epidemiological studies by Barker (1990) on *the fetal origins hypothesis*, recent scholarship has moved towards identification strategies such as family fixed-effects and natural experiments to assess the causal effect of *in utero* shocks on child development (Almond and Mazumder 2013). Natural experiments provide a random source of variation of birth endowments that is not associated with any possible confounder of the relationship of interest.

This literature studied the effect of several natural experiments[26] on birth outcomes. Despite these shocks affecting foetal health through various biological mechanisms (e.g., maternal stress, exposure to developmental toxicant, nutrition), they – from the mildest to the most severe – consistently affect birth and later developmental outcomes (Almond, Currie and Duque 2018).

However, within this literature, only a minor *corpus* of research has investigated the effect of *in utero* shocks on cognitive or socio-economic outcomes later in life (Almond 2006; Almond et al. 2009; Torche 2018, Susser and Stein 1994; van Ewijk 2011). This is most likely the case due to data limitations in finding a way to precisely and reliably link a shock occurred during the gestation stage to future outcomes. In agreement with previous findings, also these studies showed consistent effects of prenatal health shocks on various children and adult outcomes, including cognitive ability, wages, welfare dependency, and schizophrenia.

There are even fewer examples of studies that attempted to instrument parental investments or parental response in addition to the health shock. Yi, Heckman, Zhang and Conti (2015) use variation in goods prices and non-labour income as instruments for monetary investments. Other studies used educational interventions such as Head Start (Aizer and Cunha 2012), cash transfers or iodine supplementation in developing counties (for a review see Almond et al. 2017). Even though natural experiments and instrumental variables produce robust evidence on the role of *in utero* health shocks and parental investments on child development, their external validity is limited.

An ideal strategy to link *in utero* shocks with children's skill development is to rely on within-family models that compare siblings or twins born and raised in the same family. Twin fixed-effects models control for any unobserved differences across families correlated with both children's BW and later academic skills. In Table 1, we provide a comprehensive but not systematic review of the most recent and relevant literature using family fixed-effects and between-family models. To the best of our knowledge, during the last two decades, only 18 studies (only Goosby et al. 2009, and Cheadle and Goosby 2010 published in sociological journals) analysed the association between BW and academic outcomes by parental SES. Among them, only five studies jointly evaluated the effect of BW on parental investments and

[26] Terrorist attacks (Currie and Schwandt 2015), wars (Camacho 2008; Torche, 2011), criminal violence (Brown 2018; Torche and Villarreal 2014), pollutants (Almond, Edlund, and Palme 2009; Isen, Rossin-Slater and Walker 2017), natural disasters (Currie and Rossin-Slater 2013; Kim, Carruthers, and Harris 2017; Torche 2011), sporting events (Duncan, Mansour, and Rees 2017), and grievance (Black, Devereux, and Salvanes 2016; Persson and Rossin-Slater 2018).

academic outcomes (Abufhele et al. 2017; Bharadwaj et al. 2018; Cabrera-Hernández 2016; Lynch et al. 2016; Yi et al. 2015).

As shown in Table 1, seven out of 18 studies found no heterogeneity by SES in the effect of BW on educational outcomes (Abufhele et al. 2017; Bharadwaj et al. 2018; Lynch et al. 2016; Cheadle and Goosby 2010; Goosby et al. 2009; Oreopoulos et al. 2008; Black et al. 2007). Four studies only found SES heterogeneity in some specific outcomes (Curry et al. 1999), with compensatory patterns among high-SES families in three of them (Møllegaard 2020; Asbury et al. 2006; Conley et al. 2007). The remaining seven studies found fairly consistent heterogeneous effects of BW by parental SES. Among them, five found compensation (or smaller effects of BW) in high-SES families, and reinforcement (or larger effects of BW) in low-SES families (Boardman et al. 2002; Kelly et al. 2001; Lin et al. 2007; Torche and Echevarría 2011; Yi et al. 2015), while only the remaining two found reinforcement in high-SES and compensation in low-SES families (Cabrera-Hernández 2016; Figlio et al. 2014).

Overall, the reviewed literature points to the detrimental effect of LBW and twins' differences in BW on academic outcomes: a negative effect that remains fairly stable from the first grade of primary to adolescence in those very few studies combining longitudinal and twin data (Figlio et al. 2014). The patterns of heterogeneity of this effect by parental SES are mixed. Nevertheless, among those studies that found heterogeneity for at least some outcomes (11 out of 18 studies reviewed), evidence for compensation among high-SES families is more commonly found than reinforcement (8 out of 11 studies) (Boardman et al. 2002; Kelly et al. 2001; Lin et al. 2007; Torche and Echevarría 2011; Yi et al. 2015).

Regarding the association between BW and parental investments, results are also mixed. When using sibling models, some researchers find reinforcing or neutral parental response (Datar 2010), while others find compensation among high-SES families (Hsin 2012; Restrepo 2016). Among those very few studies using a twin design, parents do not respond to twins' differences in BW, allocating investments in a neutral way. Moreover, this pattern does not vary by parental SES (Abufhele et al. 2017; Bharadwaj et al. 2018). In turn, Yi et al. (2015) argue that parents compensate for prenatal health shocks with health investments but reinforce with educational investments (with low-SES parents more likely to do so).

Table 1. Literature review for studies on the effect of BW on parental investment and educational outcomes by parental SES

Year	Author(s)	Country	BW ➜ Parental Investment (I)	BW ➜ Educational Outcome (Y)	Age	Results	Results by SES	Design
1999	Curry et al.	UK	-	Test scores (O-Level in maths and English)	16	Yes	No heterogeneity Low: Compensation (Men)	Between-family
2001	Kelly et al.	England	-	Behavioural problems (SDQ)	4-15	Yes	High: Compensation Low: Reinforcement	Between-family
2002	Boardman et al.	USA	-	Maths and reading comprehension	6-14	Yes	High: Compensation Low: Reinforcement	Sibling FE
2006	Asbury et al.	England/Wales	Discipline/Negative feelings/Instructive-Informal/communication	Behavioural problems and test scores (maths and language)	7	Yes (BW➜Y)	No heterogeneity High: Compensation (BW➜Y)	MZ-Twin FE
2007	Black et al.	Norway	-	IQ; high-school completion	18-20	Yes	No heterogeneity	Twin FE
2007	Conley et al.	USA	-	Literacy, numeracy, reading comprehension, and problem-solving skills; and behaviour problems	3-12	Null	No heterogeneity Low: Reinforcement (Reading comprehension)	Sibling FE
2007	Lin et al.	Taiwan	-	Attending college	18	Yes	High: Compensation Low: Reinforcement	Between-family
2008	Oreopoulos et al.	Canada (Manitoba)	-	High-school completion	17	Yes	No heterogeneity	Twin/Sibling FE
2009	Goosby et al.	USA	-	Maths and reading comprehension growth	5-14	Yes	No heterogeneity	Sibling FE
2010	Cheadle et al.	USA	-	Test scores; high-school completion	5-19	Yes	No heterogeneity	Sibling FE
2010	Datar et al.	USA	Breast-feeding, well-baby visits, immunisations, and preschool attendance.	-	0-6	Yes	No heterogeneity	Sibling FE
2010	Torche et al.	Chile	-	Test scores	9	Yes	High: Compensation Low: Reinforcement	Twin FE
2012	Hsin	USA	Time reading, playing, hobbies, homework	-	0-12	Null	High: Compensation; Low: Reinforcement	Sibling FE
2014	Figlio et al.	USA (Florida)	-	Test scores	9-14	Yes	No heterogeneity	Twin/Sibling FE
2015	Yi et al.	China	Monetary investment in health and education	Test scores; Behavioural problems	11	Yes (BW➜I) (BW➜Y)	High: Compensation (BW➜Y) Low: Reinforcement (BW➜Y)	Twin FE
2016	Cabrera-Hernández	Mexico	Expenditure in books, fees, uniforms, tutoring	IQ; years of schooling; school attendance; grade repetition	5-17	Yes (BW➜IQ)	High: Compensation (BW➜I) Reinforcement (BW➜Y) Low: Null (BW➜I) Compensation (BW➜Y)	Sibling FE
2016	Grätz et al.	USA	Cognitive stimulation at 2 years-old	-	2	Null	No heterogeneity	Twin FE
2016	Lynch et al.	USA	Parental interactions; investments; warmth; parenting quality; cognitive stimulation	Test scores	5	Yes (BW➜I) (BW➜Y)	No heterogeneity	Between-family
2016	Restrepo	USA	HOME	-	0-14	Null	High: Compensation Low: Reinforcement	Sibling FE
2017	Abufhele et al.	Chile	HOME; Maternal time investment	Test scores: Cognitive, Motor, Language, Socio-emotional	0-7	Null (BW➜I) Yes (BW➜Y)	No heterogeneity	Twin FE
2018	Bharadwaj et al.	Chile	Maternal time investment in educational activities	School grades and test scores	6-14	Null (BW➜I) Yes (BW➜Y)	No heterogeneity	Twin/Sibling FE
2020	Møllegaard	Denmark	-	Behavioural Problems (SDQ)	12	Yes	High: Compensation Low: Reinforcement	MZ-Twin FE

4. Data, Variables and Sample Selection

4.1. Data

We draw data from the first wave of the *Twin Life Study – Genetic and Social Causes of Life Chances*, a cross-sequential panel study comprising four age cohorts of same-sex twins aged 5 (born 2009/10), 11 (born 2003/04), 17 (born 1997/98), and 23-24 (born 1990/93) (Diewald et al., 2018). We analyse the youngest birth cohorts (n=4,096) of same-sex MZ and DZ twins at ages 5 (born 2009/2010) and 11 (born 2003/2004). The *Twin Life Study* was designed as a probability-based sample intended to be representative of German municipalities and rural areas, and families with same-sex twins (Brix et al., 2017). The selected municipalities provided a random sampling of twin families within the specified age groups of the twins. The sample was drawn from administrative registries of residents by identifying those individuals with identical addresses, birthdays and genders. The first face-to-face wave of the study was carried out between 2014 and 2016, interviewing twins, siblings and parents with CAPI, CASI and PAPI survey methods, with a participation rate of about 40%. The second face-to-face wave was collected between 2016 and 2018 and was published in the spring of 2020. However, none of the outcome variables that we study in the first wave were available in the second wave. This limitation, in addition to attrition rates at about 30% in the second wave (Lessar et al., 2020), prevented us from carrying out longitudinal analyses within age cohorts.

Technical reports of the *Twin Life Study* compared distributions of the key socio-demographic variables of the survey with the German micro-census survey by identifying a proxy-twin and a multiple-child household sample (Lang and Kottwitz 2017). This report concluded that (1) (proxy-)twin and multiple-child households (3 or more) identified in the micro-census are comparable in different socio-demographic characteristics; and (2) that the *Twin Life* sample is relatively comparable to German households with multiple children, covering the full distributions (lower and upper-bounds–)of the parental SES variables. Still, the sample of the *Twin Life Study* is positively selected in terms of urban households, German citizenship, parental socioeconomic status and mothers' age, and very negatively selected when it comes to BW and prematurity. Hence, twin fixed-effects estimates are upper-bound estimates of BW, especially when analysing MZ twins that share a single placenta, even when twins are more closely monitored during pregnancy, twin-FE suffer from attenuation bias due to measurement error, and high-SES families tend to be overrepresented. In the Appendix, Table A.16. illustrates the selection bias of twins by showing a comparison of summary statistics of some key variables between the *Twin Life* sample and a representative sample of the German population of children at school entry drawn from the *National Educational Panel Study*. Also, Figure A.2. in the Appendix illustrates the distribution of BW among twins and singletons.

4.2. Variables

Birth weight. Birth weight information is collected from scanned health book records for 83% of the sample and from parental reports for the remaining 17%[27]. Birth weight is measured on a continuous scale in grams. We transformed the original scale into logarithmic to account for skewness and nonlinearities in its effect. Given that gestational age varies across families and affects BW, we also operationalise BW by adjusting for gestational week as a robustness check (Torche and Echevarria 2011:1010)[28]. This measure produces consistent results with the logarithm of BW in the SES heterogeneity analyses (see Appendix Table A.7.). There are alternative measures of health at birth available in the health book records such as birth length and head circumference. Unfortunately, the large share of missing values of these variables with respect to BW prevents us from using them.

Academic performance. For the 5-year-old cohort, no information on school grades is yet available. Thus, we use a standardised average of parental ratings on two academic competencies as a proxy for academic performance: (i) language skills in German; and (ii) mathematical skills. Originally, parents are asked to compare their child's degree of similarity in these competencies with other kids outside the family in a scale ranging from 1 ('a lot worse') to 5 ('a lot better'). For the 11-year-old cohort, data on academic performance is collected by taking photos of the most recent report card (grades 4-6)[29] of the children's school grades in mathematics and German. Originally, grades vary from 1 (excellent) to 6 (insufficient). We reversed the scale and created a standardised average of school grades in mathematics and German. Measure comparability by cohorts may be an issue, hence, in Appendix Figure A.1. we show that possible reporting bias[30] in the 5-year-old cohort does not

[27] Retrospective reports of birth weight may induce recall bias. However, previous research shows that recall bias is slight even 15 years after the birth, with discrepancies between parental reports and hospital records between 10 and 50 grams. Moreover, misreporting does not seem to vary by parental SES (O'Sullivan et al. 2000; Yawn et al. 1998).

[28] Following the strategy by Torche and Echevarria (2011:1010), we estimate "the difference between individual birthweight and gestational week-specific mean birthweight, dividing by the gestational week-specific standard deviation of birthweight", and standardising and adjusting by sex.

[29] Even though school grades were taken from different tracks (primary, vocational training or academic track between grades 4 and 6), the twins' intra-class correlation (ICC) stands at 0.95, and we control for this variable in the models predicting academic performance for the 11-year-old cohort. Furthermore, a restricted analysis of only twins attending grade 4 of primary education yields equivalent results to the main analysis.

[30] Measurement error is likely an issue for between-family comparisons when using parental ratings of children's skills. For fixed-effects and within-family models, though, this bias should not be a major concern. Parental ratings are subjected to *reference bias.* This is a classic problem in measuring subjective health or personality traits. The reference group for comparisons –*children outside the family at the same age*– of low-

seem related to parental SES, as the SES-gap in academic performance remains constant over cohorts.

Cognitive ability. Cognitive ability is measured with the *Culture Fair Test*, a widely used and well-validated cognitive test battery that captures non-verbal (fluid) intelligence as a proxy for general cognitive abilities. For the 5-year-old cohort, the version *CFT 1-R* of the *Culture Fair Test* is used, comprising three subtests on figural reasoning (15 items), figural classification (15 items), and matrices (15 items). The test is administered in a paper-pencil format by trained interviewers. For the 11-year-old cohort, the version *CFT 20-R* of the *Culture Fair Test* was administered, comprising four subtests on figural reasoning (15 items), figural classification (15 items), matrices (15 items), and reasoning (topology) (11 items). The test is administered via computer, resulting in a sum of all correctly answered items in a battery of four subtests. We applied a latent factor approach to construct a standardised cognitive ability score from the items for each cohort independently, with satisfactory Cronbach's alpha at 0.72 (5-year-old cohort) and 0.70 (11-year-old cohort).

Behavioural problems. Behavioural problems are measured with the adapted *Strengths and Difficulties Questionnaire* (SDQ) for the age range 3–16. The SDQ is a brief emotional and behavioural screening questionnaire for children and adolescents widely used in psychiatry and psychology to assess mental health, behavioural problems and positive child development. For the 5-year-old cohort, behavioural problems are reported by the parents, while for the 11-year-old cohort they are self-reported by the children. SDQ comprises twenty subdomains on externalising (10 subscales on hyperactivity/attention and behaviour problems) and internalising problems (10 subscales on emotional symptoms and social problems) ranging from 0 to 2, where 0=*does not apply at all*; 1=*partly applies*; and 2=*applies completely*. We computed the total score of behavioural problems as the sum of the twenty subdomains, with satisfactory Cronbach's alpha at 0.78 (5-year-old cohort) and 0.76 (11-year-old cohort). From the total sum of the index, we standardised the scores.

Parental SES. Parental SES is measured with a dummy on the highest educational level (ISCED-97) achieved by the parents: 0=ISCED 1-4 and 5B (<university) and 1=ISCED 5A-

---

and high-SES families may differ depending on residential segregation, sociability, and kindergarten attendance/quality (see section 6.3.3. labelled "Additional Mechanisms: School Investments" and Appendix Tables A.13.–A.14. for a detailed analysis). The correlation between objectives measures of children's cognitive skills and parental reports are higher for high-SES families (and their ICC is lower), suggesting that they are more able to rate the "true ability" of their kids than low-SES parents. Thus, if anything, we would expect more measurement error among low-SES families and, consequently, underestimation/attenuation of coefficients in fixed-effects models.

6 (university of applied sciences, university and Ph.D.). The main reason to codify this variable in such a reduced way is to maximise sample size to split the analysis by parental SES. With this categorisation, the analytic sample size is balanced with around 50% of total cases in each SES subgroup. Parental educational attainment is considered as a good proxy for cultural resources of the families and it is a good predictor of children's educational performance (Jæger 2011). Moreover, as parental SES measures are measured way after birth, parental education is the most reliable and time-stable SES indicator to capture prenatal and postnatal family environment. Nonetheless, the measurement of parental SES via parental education may be problematic under the classical human capital framework. Most of these theories (be it of compensation, reinforcement or between families) deal with the allocation of limited resources —economical resources and time—and education may not be the best measure of these, capturing other relevant factors that are not in limited supply. Thus, we also carried out sensitivity analyses with an alternative measure of parental SES using the highest parental *International Socio-Economic Index of Occupational Status* (ISEI) codified into a dummy (0=q1-q2, and 1=q3-q4). Results are robust to this alternative specification (see Appendix Tables A.3.–A.4. below).

Parental investments. We use two key measures of parental involvement for child development: parental time (e.g., weekly frequency) in cultural activities, and parental emotional warmth. These measures are twin-specific and observed just before the survey. We have to plausibly assume that these measures are a good proxy for the accumulated record and absolute level of investments across families. However, we acknowledge four limitations. First, these measures might not fairly represent the past (unobserved) patterns of allocation of investments within families. Second, measures of parental investments are observed in an unspecified retrospective window before the survey. Thus, endogeneity could be an issue when assessing the moderating effect of parental investments on academic outcomes, as the former might be responsive to previous skills. Third, twins' comparisons of investments are subjected to parental inequity aversion[31], *common-goods* and *spillover* effects, making it more difficult to detect reinforcing parental response. However, as can be seen in Table 2 and Figure 4 below, there is considerable within-family variation in parental allocation of investments to be analysed (0.75-0.82 SD average difference between twins). Mönkediek et al. (2020) tested whether parents treat twins more similarly than non-twin siblings and they concluded that

[31] Furthermore, parents may be especially reluctant to report differences in investments between twins in comparison to children's reports (Bharadwaj et al. 2018:358).

twins do not receive more differential treatment than non-twins once age differences are controlled for. They argue that these findings make twin studies generalisable to non-twin families. Fourth, parenting measures display low reliability, and within-family variation may contain a large part of random noise. Unfortunately, self-reported ratings of parenting by children and/or parents are the most common measures used in this field of research to assess the role of parental involvement. Only in very detailed studies conducted by developmental psychologists is there information available on direct observation of parent-child relationships, raising also concerns on desirability bias when behaviour is being recorded.

Parental involvement in cultural activities. Parental involvement in cultural activities is measured by the parental (5-year-old cohort) and child's self-report (11-year-old cohort) of the frequency of activities with family members in the last four weeks. This measure comprises four (5-year-old cohort: i-iv) or five items (11-year-old cohort: i-v) on: (i) singing/making music; (ii) sports; (iii) reading/talking about books/stories; (iv) activities with family members: playground, walks, day trips; and (v) visits to museums/theatres. These items range from 1 (*not at all*) to 5 (*almost daily*). From these items we computed a standardised index of parental involvement in cultural activities for each birth cohort, with a Cronbach's alpha at 0.54 for the 5-year-old cohort and 0.66 for the 11-year-old cohort.

Parental warmth. For both birth cohorts, parental warmth is measured by children's reports on four items on the parental frequency (e.g., an average of responses to independent questions on mother and father) of *showing affection*, *praising*, *cheering up*, and *supporting*. These items range from 1 (*never*) to 3 (*very often*) in the 5-year-old cohort, and from 1 (*never*) to 5 (*very often*) in the 11-year-old cohort. From these items we computed a standardised index of parental warmth for each birth cohort, with a Cronbach's alpha at 0.60 (5-year-old cohort) and 0.76 (11-year-old cohort). It could be argued that a parenting practice/behaviour such as warmth may not lend itself very well to a formalised human capital framework wherein resource constraints play a central role in influencing parental investments in the same way as money or time. However, there is ample evidence on the association between parental SES, exposure to stress, and related psychological problems that negatively affect parenting quality in terms of warmth. Thus, we argue that warmth can be seen as a child investment—in the sense that it is associated with child skill formation—that might carry an emotional cost (love and emotional responsiveness are neither countable nor infinite, especially after a long working day) and parents might consciously or unconsciously distribute unequally among their children. Of course, many parents will have trouble recognising or admitting to this behaviour, but it is not far-fetched to say that parents tend to have preferences or favouritisms among their children. The bottom line is that parents may be more attached and emotionally close to one of their twins (e.g., the lighter or heavier one in terms of BW), while also loving and caring for the other twin.

4.3. Sample Selection

In Table 2, we show descriptive statistics of each variable by birth cohort (see Table A.1. for descriptive statistics by SES). Table 3 displays the missing and excluded cases from the overall sample by birth cohort and parental SES. For the 5-year-old, the share of missing values is low (below 8%) for all variables except for cognitive ability, with 25% missing cases. For the 11-year-old cohort, the share of missing values is considerable in the parenting variables (around 18%), and academic performance (20%). Generally, in both birth cohorts, the incidence of missing information is slightly larger for low-educated families. This means that, if anything, socioeconomic inequalities might be underestimated. We use different analytic samples for each outcome and parenting variable to maximise sample size instead of applying list-wise deletion.[32] The analytic samples range from 1,396 to 1,802 observations for the 5-year-olds cohort, and from 1,262 to 1,826 observations for the 11-year-old cohort. We split these samples by parental SES. In Appendix Table A.2., we carry out a sensitivity analysis predicting missing data for each analytic sample of the 11-year-olds cohort to conclude that the sample characteristics do not vary systematically. We only find that the analytic samples for parental warmth are slightly more positively selected in terms of SES.

[32] We did not apply multiple imputation to solve the problem of small sample size, missing data, and attenuation bias in twin fixed-effects because there are not good auxiliary variables that vary within families and might predict missing outcome variables.

Table 2. Descriptive statistics by birth cohort

	5-year-old cohort							11-year-old cohort						
Variables	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**ICC**	**Alpha**	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**ICC**	**Alpha**
Female twin-pair	2,014	0.51		0	1			2,082	0.52		0	1		
Dizygotic twin-pair	2,014	0.57		0	1			2,082	0.60		0	1		
Birth order	2,014	1.5		1	2			2,082	1.5		1	2		
Age in months	2,014	65.44	3.87	52	75			2,082	137.43	3.74	130	146		
Mother's age	2,006	37.28	5.13	22	59			2,072	43.01	4.89	28	58		
High Parental Education	1,998	0.44		0	1			2,070	0.46		0	1		
Highest Parental ISEI	1,908	54.07	27.71	*0	89			1,944	53.18	26.53	0	89		
High Parental ISEI (q3-q4)	1,908	0.51		0	1			1,944	0.48		0	1		
Parental involvement in cultural activities (raw)	1,960	2.40	0.38	1.03	3.26	0.97	0.54	1,688	1.53	0.53	0.66	3.30	0.59	0.66
z - Parental involvement in cultural activities	1,960	0.00	1.00	-3.67	2.31	0.97	0.54	1,688	0.01	1.00	-1.64	3.64	0.57	0.66
z -Within-family absolute differences in cultural activities	1,960	0.17	0.30	0	3.27			1,688	0.82	0.72	0	4.41		
Parental warmth (raw)	1,990	4.56	0.37	3	5	0.91	0.60	1,728	4.29	0.66	1.5	5	0.63	0.76
z - Parental warmth	1,990	0.00	1.00	-4.2	1.19	0.91	0.60	1,728	-0.11	1.05	-4.56	1.03	0.63	0.76
z -Within-family absolute differences in parental warmth	1,990	0.33	0.46	0	5.35			1,728	0.75	0.70	0	4.59		
Birth weight (grams)	1,852	2,353.24	557.09	280	3,880	0.89		1,896	2,420.12	538.97	630	4,300	0.87	
Within-family birth weight differences (grams)	1,852	272.97	231.85	0	1,920			1,896	269.97	250.58	0	1,800		
Log(birth weight)	1,852	7.73	0.28	5.63	8.26			1,896	7.76	0.26	6.45	8.37		
LBW	1,852	0.57		0	1	0.70		1,896	0.51		0	1	0.71	
z – Foetal growth	1,852	0.00	1	-5.20	2.95	0.59		1,896	0.00	1	-3.84	4.91	0.61	
Gestation week	1,852	35.60	2.70	22	44			1,896	35.85	2.55	24	42		
Preterm	1,852	0.55		0	1			1,896	0.53		0	1		
Physical/mental illness diagnosis	1,978	0.47		0	1	0.69		2,034	0.56		0	1	0.64	
z -Cognitive ability	1,518	0.05	1.00	-2.46	3.53	0.78	0.72	1,896	0.02	0.99	-3.36	2.86	0.67	0.70
z -Academic performance	1,970	0.00	1.00	-3.14	2.34	0.73	0.60	1,548	0.02	0.99	-3.23	2	0.76	0.69
Behavioural problems (raw)	1,956	7.54	4.64	0	30	0.66	0.78	2,014	10.19	5.19	0	30	0.51	0.76
z -Behavioural problems	1,956	0.00	1.00	1.47	5.90	0.64	0.78	2,014	0.00	1.00	-1.78	4.12	0.53	0.76

Notes: ICC=intra-class correlation; Alpha=Cronbach's alpha.; *0=unemployed.

Table 3. Sample selection and missing data by birth cohort and parental education

Variables	5-year-old cohort						11-year-old cohort					
	All sample		Low Education		High Education		All sample		Low Education		High Education	
	n missing	% missing	n missing	% missing	n missing	% missing	n missing	% missing	n missing	% missing	n missing	% missing
Highest parental education	16	0.79%					12	0.58%				
Highest parental ISEI	106	5.26%					138	6.63%				
Parental involvement in cultural activities	54	2.68%	32	2.86%	18	2.05%	394	18.92%	210	18.95%	182	18.92%
Parental warmth	24	1.19%	16	1.43%	8	0.90%	354	17.00%	224	20.22%	124	12.89%
Birth weight	162	8.04%	98	8.77%	60	6.82%	186	8.93%	104	9.39%	80	8.32%
Cognitive ability	496	24.63%	288	25.76%	198	22.50%	186	8.93%	108	9.75%	78	8.11%
Academic performance	44	2.18%	22	1.97%	22	2.50%	414	19.88%	236	21.30%	176	18.30%
Behavioural problems	58	2.88%	26	2.33%	30	3.41%	68	3.27%	28	2.53%	40	4.16%
n	2,014		1,118		880		2,082		1,108		962	

Notes: including non-missing cases within unbalanced twin-pairs as missing.

5. Empirical Strategy

5.1. Identification Strategy: Twins as a Natural Experiment

Birth weight has two main determinants: gestational length (namely the time a child stays in the mother's womb with an average of 40 weeks) and intrauterine foetal growth (the gestational age-specific growth of the foetus) (Torche and Conley 2016). Both determinants have different underlying biological mechanisms and aetiologies. For instance, hormonal fluctuation may shorten the gestation and consequently result in premature and underweight deliveries (McLean et al. 1995). Similarly, changes in the maternal-placental blood exchange and oxidative stress reduce the flow of nutrients and oxygen, resulting in reduced foetal growth (Slama et al. 2008). The contribution of these two determinants to the prevalence of LBW deliveries in Western societies differs, with about 70% of LBW kids being premature but normal for gestational age (Torche and Conley 2016).

Beyond the biological causes of BW, also social factors play a crucial role in shaping birth outcomes. There are BW inequalities among the classic racial and socioeconomic divides. Low-educated mothers are two times more likely to deliver a LBW child than high-educated mothers (Currie 2011). There are many potential factors explaining why disadvantaged mothers are less able to provide a healthy foetal environment for their children. For example, stratified unhealthy prenatal behaviour such as smoking accounts for BW differences across the socioeconomic strata (Härkönen et al. 2018; Pampel, Krueger and Denney 2010; Raisanen et al. 2013). Poor mothers are also more likely to be under chronic stress in their everyday life, which in turn might have a detrimental effect on BW (Aizer, Stroud and Buka 2016; Torche 2011). Finally, disadvantaged mothers are more likely to experience residential segregation, being exposed to high level of pollutants and harmful chemicals (Currie 2011; Slama et al. 2008).

An ideal test of the *compensatory advantage hypothesis* would compare individuals who differ in nothing but their birth endowments. Nature provides an experimental setting with the incidence of twins (Knopik et al. 2017). Twin comparisons rule out most sources of variation between and within families (e.g., birth spacing and order, mother's age, sibling-specific shocks) that might confound the association between BW and later educational outcomes (Currie 2011). Twin fixed-effects models control for most possible sources of biological and social inheritance of BW: parental body mass index (BMI), prenatal investments and behaviour (e.g., smoking, diet, antenatal health care), maternal age, gestational age, at least 50% of genetic differences, and parental SES.

Differently from singleton births, in multiple births, twins share at least 50% of their genes, the same prenatal environment and gestational length. Additionally, our *Twin Life* sample only includes same-sex twin pairs.[33] Consequently, among twins, the only source of variation in BW is due to twin differences in intrauterine foetal growth. Previous studies applying causal inference methods consider variation in twins' BW as an identification strategy as good as random (Torche and Conley 2016). This assumption is drawn from the medical literature, considering twin-differences in BW being due to the following factors (see Cleary-Goldman and Alton, 2008 for a systematic review of the aetiology of growth abnormalities in multiple gestations): (1) structural abnormalities (placenta and/or umbilical cord); (2) adverse intrauterine factors, such as small placental weight, single umbilical artery, excessive velamentous cord insertions, infections (Victoria, Mora and Arias, 2001), or various placental abnormalities; and (3) twins' *competition* for nutrients, oxygen and space (twin foetus-specific position) *in utero* from the third trimester of gestation, especially in the case of MZ twins sharing a single placenta (monochorionic twins, representing around 75% of MZ-twins pregnancies). These factors are assumed to affect randomly only one twin in a given pair but, this is, of course, just an unstable assumption. Further evidence on this issue is necessary to assess the internal validity of twin-discrepancies in BW as a random natural experiment.

Among DZ-twins, genetic differences also account for a small portion of variation in BW (Gielen et al. 2008). Genetic differences in BW among DZ twins born in the same family are a product of random segregation of alleles in the process of reproduction. However, it could be the case that some of the genetic variants that explain BW also contribute to cognitive/non-cognitive development, inducing unobserved confounding. However, according to previous research, the confounding potential of common genetic correlation between BW and skills is minor (Newcombe et al. 2008; Conley et al. 2018). We carried out a robustness check of the main analysis stratifying by twins' zygosity[34] (see Appendix Table A.6.) to conclude that the causal effect of BW is not compromised by genetic confounding, as

[33] In Appendix Table A.6. we carried out a heterogeneity analysis by twin-pair gender to conclude that, at age 11, the effect of BW on cognitive ability and behavioural problems is more detrimental for females (Møllegaard 2020). At age 5, the effect of BW on academic performance is more detrimental for males.
[34] Fortunately, Mönkediek et al. (2020) tested if parents treat MZ twins more similarly than DZ twins and they concluded that the role of twin zygosity in explaining differential parenting, in terms of emotional warmth, is very limited.

we consistently find a causal effect of BW among MZ-twins, even larger than that for DZ-twins.

Applying classic behavioural genetics methods of variance decomposition in BW to the *Twin Life* sample (detailed results available upon request) (Knopik et al. 2017), total variance in BW is explained by shared environmental factors that vary between families (73.1%), genetic differences between and within families (DZ twins) (4.5%), and twin-specific environmental factors within families (22%). When estimating twin-FE models, we exploit the last two sources of variation. Hence, twin pairs discordant in BW can be thought of as a natural counterfactual in which the co-twins can be used as their own control/experimental group (McGue et al. 2010).

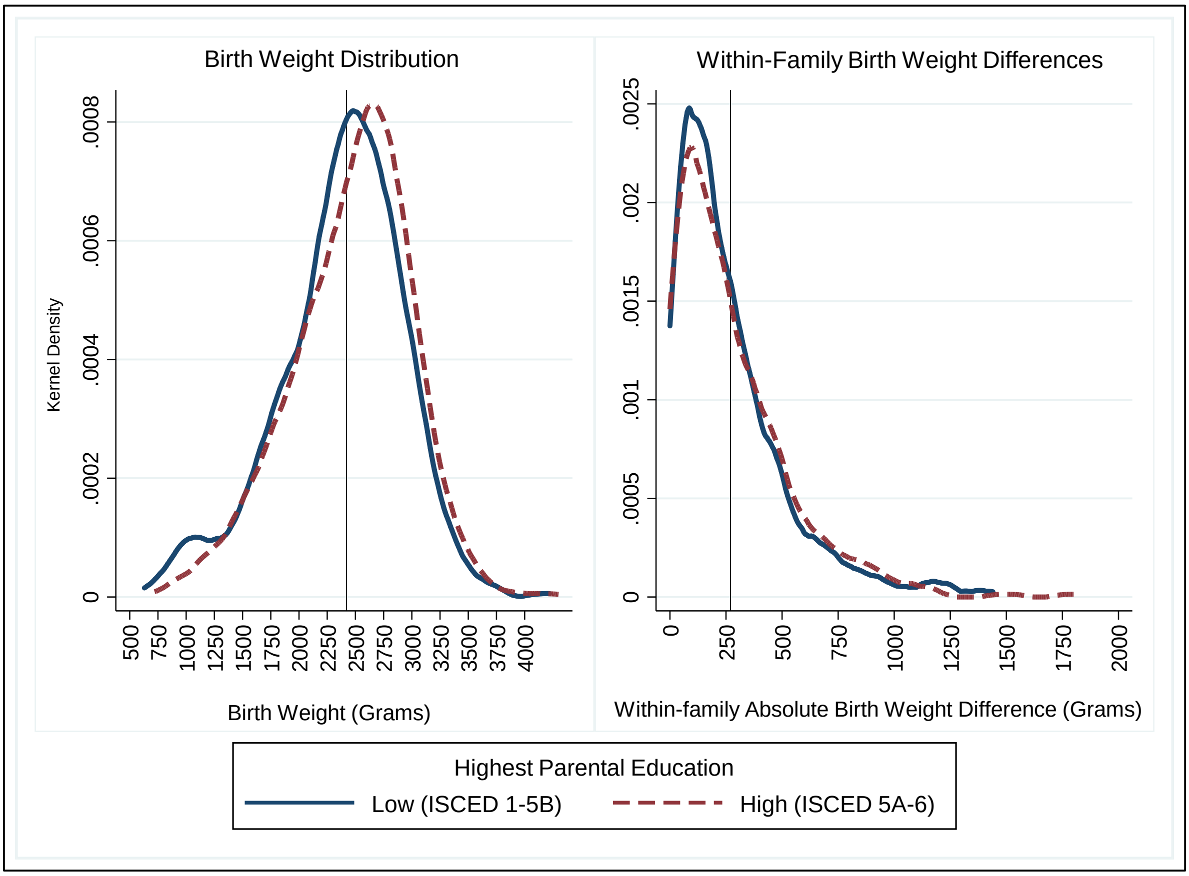


Figure 3. Distribution of BW and within-family BW differences by parental education for the 11-year-old cohort

Figure 3 illustrates the variation in BW across families (left-hand graph) and within families (right-hand graph) by parental SES. In the empirical analysis, we exploit variation in BW within-families, with an average twin difference in BW at 273 grams (5-year-old cohort) and 269 grams (11-year-old cohort). Importantly, as can be seen in Figure 3 and in Appendix Table A.1., average within-family differences in BW are virtually the same for low (275 grams – 5-year-old cohort; 267 grams – 11-year-old cohort) and highly-educated families (269 grams – 5-year-old cohort; 273 grams – 11-year-old cohort).

However, it still could be argued that within-family differences in BW are drawn from absolute BW distributions that differ by SES. As shown in left-hand Figure 3, low-SES families have a larger prevalence of LBW (≈6%) that may capture SES-differences in prenatal behaviour and/or access to and quality of antenatal care. Thus, within-family differences in BW may be more detrimental for low-SES than high-SES families not due to SES-gaps in postnatal investments, but because the former are more likely to be drawn from the LBW risk group. To account for this possibility, in Appendix Table A.5., we carry out a robustness check across the absolute BW distribution and parental SES among families under or above the LBW threshold (see *section 6.3.1* below). Additionally, we assess prenatal parental behaviour and antenatal care by SES (see *section 6.3.1* below).

Finally, in mediation and moderation analyses, we use measures of parental investments that are not exogenous, as they are assessed after BW (Aizer and Cunha 2012). Hence, there might be unobserved factors (i.e., parental choice) explaining both twins' differences in allocation of investments and academic outcomes, being very complex to predict *a priori* the direction of the potential bias. However, we are not aware of any suitable instrument to randomise parental investments available in our data, with external validity, or that would meet the exclusion criteria (Grätz and Torche 2016:1891). Figure 4 illustrates the distribution of within-family differences in parental investments.

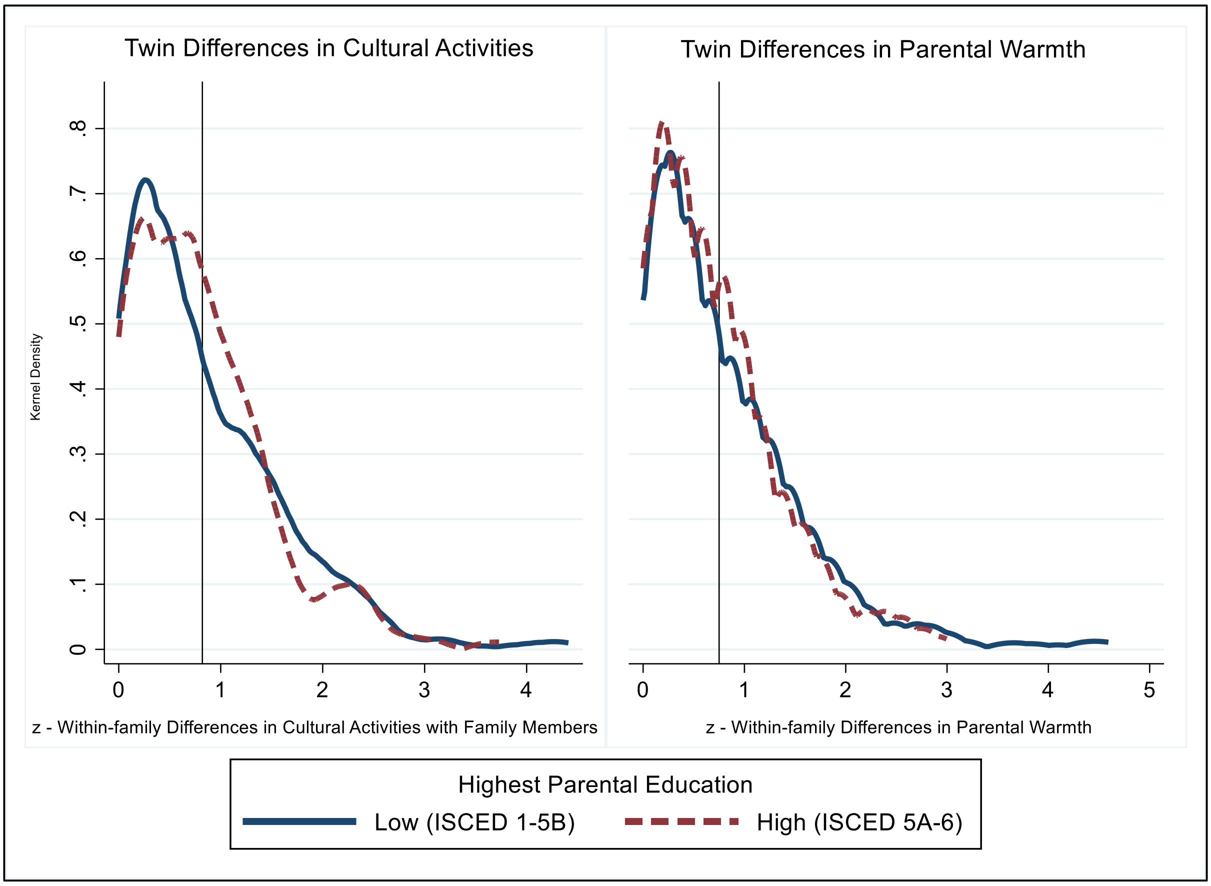


Figure 4. Distribution of within-family differences in parental investments among twins by parental education for the 11-year-old cohort

5.2. Twin-Fixed-Effects Estimators

The paired structure of the data comprises two twins[35] (j = first-born twin; k = second-born twin) clustered in families (i), keeping only balanced pairs. We estimate OLS twin-fixed-effects (FE) models[36] with clustered standard errors at the family level to account for unobserved correlation of twins within-families.

Equation 1a shows the baseline twin-FE model in which Y is the skill outcome; β_1 stands for the main coefficient of interest on the causal effect of the natural logarithm of BW

[35] The presence of additional siblings, which it is the case in about 50% of the families for low- and high-SES families alike, may influence intra-family allocation decisions, making it (even) more complex to test and interpret intra-family allocation theories, hypotheses, and empirical estimations. The mean age-difference between twins and the non-twin sibling is about 3 years, being full siblings older on average with a minimum age of 5 years. Thus, resource constraints might be especially salient for low-SES families when allocating their resources and investments between 3 (or more) children (e.g., *resource dilution*). Unfortunately, the share of families with additional siblings further than twins is around 50%, preventing us from carrying out a further heterogeneity analysis by household structure, as the sample is already quite small after applying list-wise deletion and estimating models independently by parental SES. Moreover, please keep in mind that the main fixed-effects models do not allow to control for family-constant variables such as number of siblings.

[36] As a robustness check, in Appendix Table A.10. we also run Naïve OLS models treating twins as individual observations with clustered standard errors and controlling for between-family confounders. The true causal effect of BW should be found somewhere in between Naïve and twin-FE models, as the former are subjected to between-family confounders and the latter to attenuation bias.

differences (10% difference in BW, which corresponds to around 240 grams) between twin j and twin k in family i among discordant twin pairs; and $\boldsymbol{Z}$ represents a vector of covariates (birth order[37] in all models; and grades' track for the models predicting academic performance in the 11 year-old cohort). One should note that the causal biological effect of BW may be overstated if parents invest more in heavier co-twins (reinforcement) or underestimated if parents allocate more investments in lighter co-twins (compensation) (Yi et al. 2014). On the right-hand side of equation 1a, e_{ij} and e_{ik} stands for the twin-specific error of prediction in family i. Equation 1a is estimated by birth cohort for all the sample and independently by parental SES. Equation 1a can also be expressed more succinctly as a difference operator in equation 1b.

$$Y_{ij} - Y_{ik} = \beta_1(logBW_{ij} - logBW_{ik}) + \beta_2(\boldsymbol{Z}_{ij} - \boldsymbol{Z}_{ik}) + (e_{ij} - e_{ik}) \ (1a)$$

$$\Delta Y_{ij} = \beta_1 \Delta logBW_{ij} + \beta_2 \Delta \boldsymbol{Z}_{ij} + \Delta e_{ij} \ (1b)$$

Even when twin-FE models provide solid control for most possible omitted variables, two main limitations should be taken into account when interpreting empirical analyses. First, as twins share the family environment and at least 50% of their genetic makeup, the amount of variation to be explained is generally small. For the outcomes studied here, as shown in Table 2, the intra-class correlation (ICC) stands between 0.70 and 0.76, meaning that 70%-to-76% of the variation in cognitive and non-cognitive skills is explained between families, and 30%-to-24% within families—that seems large enough.

Second, by estimating twin-FE we deal with limited power to find statistically significant effects, as half of the degrees of freedom are lost (Conley et al. 2019). Hence, effects should be two times larger than models using all available individuals (e.g., naïve models with clustered errors), which also provide inflated standard errors due to the non-independence of

[37] Birth order captures the order in which the co-twins were delivered in the multiple pregnancy. It is important to control for birth order within the twin pair because we intend to capture random variation in twin-specific access to nutritional intake in the pre-natal uterine environment leading to BW discordances. Birth order is itself positively correlated to BW (Yokohama et al., 2016), and later BMI and status-attainment. Birth order shares some of its causes with the determinants of twin-differences in intrauterine growth and BW (e.g., placental weight, central insertion of the umbilical cord, and twin-specific placental position), but its ultimate causes are unknown. Thus, birth order might be a confounder of the association between BW, perinatal health, and later skill formation, since it may capture different (and non-random) determinants from the ones causing BW, which are our focus. It has been found that first-born twins are slightly heavier than second-born twins (Yokohama et al., 2016) and, independently of the mode of delivery and net of BW, second-born twins have lower a BMI and a higher risk of neonatal morbidity and mortality than first-borns (Kim et al., 2020). Finally, for comparability reasons with previous research estimates, controlling for birth order is common practise in the field to assess the effect of BW. Therefore, for the sake of comparability, we also followed this approach.

observations. In addition to this issue, classical measurement error may further attenuate effects in combination with twin-FE, reducing the signal-to-noise ratio (Kohler et al. 2011). For these reasons, we also report and interpret findings with an alpha level at 10%.

Finally, following the logic of experimental design, we carried out a simulated power calculation to answer the following question: "given a sample size, how large would the true effect size have to be in order to be able to detect it with reliable power using a test of size alpha=0.05? (Currie and Almond 2011:1333-1336)." With a sample size of around 1,400 observations (smallest n in the main models), we should be able to detect a coefficient from about 0.15 at alpha=0.05 if it were a true effect.

5.3. Mediation and Moderation Analysis

We estimate partial mediation analysis to test whether and how parents allocate investments (*I*) among twins as a function of birth endowments. Equation 2 estimates the effect of BW (β_1) on parental allocation of investments or parental response ($I_{ij} - I_{ik}$). Equation 2 is estimated independently by parental SES. To find consistent mediation, there must be a link between twin differences in BW and parental investments in equation 2. We assume no heterogeneity in BW returns by parental investments within-families.

$$I_{ij} - I_{ik} = \beta_1(logBW_{ij} - logBW_{ik}) + \beta_2(\boldsymbol{Z}_{ij} - \boldsymbol{Z}_{ik}) + (e_{ij} - e_{ik}) \quad (2)$$

Finally, to test the moderating role of the absolute level of parental investments or resources across families, we estimate equation 3 with an interaction term between twin differences in BW (β_1) and I_i. I_i stands for the average of parental investments between twin *j* and twin *k* in family *i*, categorised in quintiles to account for nonlinearities in the functional form.

$$Y_{ij} - Y_{ik} = \beta_1(logBW_{ij} - logBW_{ik}) + \beta_2[(logBW_{ij} - logBW_{ik}) * (I_i)] + \beta_3(\boldsymbol{Z}_{ij} - \boldsymbol{Z}_{ik}) + (e_{ij} - e_{ik}) \quad (3)$$

In order to account for potential between-family confounding in estimating this interaction —LBW, parental prenatal behaviour, and SES—we explored the characteristics of the sample by quintiles of investments (see *section 6.3.1* and Table A.15.). Additionally, we also ran models with controls for within-family differences in investments and LBW discordancy and the results held.

6. Findings

Table 4 summarises the main results of the article from OLS twin-FE models. Models are estimated by birth cohort for all samples and independently for low and high-educated families. On the left-panel of Table 4, we can see the effect of log BW on three academic outcomes for the 5-year-old cohort, while on the right-hand side of the table the results for the 11-year-old cohort are displayed.

At 5 years old, in both low- and high-SES families, we find an effect of BW on academic performance and behavioural problems. A 10% difference in BW between twins raised in the same family (≈240 grams) is associated with 6% standard deviation (SD) units better academic performance, and with 7% SD units fewer behavioural problems. For cognitive ability, we find null effects, a puzzling result as, as we discussed above, LBW is found to be neurologically detrimental for cognitive development. One possible reason explaining this null result is the considerable share of missing values (25%) for this variable, so incurring in positive selection in cognitive ability and BW among respondents.

At 11 years old, we find again an effect of BW on academic performance and behavioural problems of a similar magnitude than in the 5-year-old cohort, but only for low-educated families.[38] Specifically, we find that a 10% difference in BW is associated with an increase by 10% SD units in academic performance, and with a decrease by 6% SD units in behavioural problems among low-educated families. For cognitive ability, we still find a null effect for both low and high-educated families. However, as shown in Appendix Table A.5., when we estimate these models at the bottom of the BW distribution (LBW; < 2,500 grams), there is an effect of BW on cognitive ability only for low-educated families (**β**=6% SD; *p-value*=10%). This result suggests that BW may be only detrimental for cognitive development under circumstances of substantial disadvantage (e.g., LBW threshold and low-SES) (Aarnoudse-moens et al. 2009).

[38] Disentangling the effect of academic performance and behavioural problems, we find a stronger effect on mathematics and externalising behaviour (attention problems).

Table 4. Twin-fixed effects (FE) OLS models for the effect of log(BW) on educational outcomes by birth cohort and parental education

	5-year-old cohort				11-year-old cohort			
Academic Outcomes	All Sample	Low-Edu.	High-Edu.	Low - High	All Sample	Low-Edu.	High-Edu.	Low - High
z - Academic performance	0.65***	0.67**	0.61*	0.10	0.74***	1.04***	0.46	0.57
	(0.21)	(0.27)	(0.34)	(0.43)	(0.22)	(0.34)	(0.28)	(0.43)
n	1,802	1,002	800	1,802	1,412	720	692	1,412
z - Cognitive abilities	0.05	-0.12	0.32	-0.42	0.17	0.41	-0.16	0.54
	(0.16)	(0.20)	(0.30)	(0.35)	(0.22)	(0.31)	(0.35)	(0.43)
n	1,396	760	636	1,396	1,716	902	814	1,716
z - Behavioural problems	-0.66***	-0.60***	-0.79**	0.24	-0.39*	-0.64**	-0.02	0.65
	(0.19)	(0.22)	(0.30)	(0.42)	(0.23)	(0.31)	(0.35)	(0.46)
n	1,792	998	794	1,792	1,826	982	844	1,826

Notes: FE=OLS fixed effects; robust standard errors in parentheses; the coefficients multiplied by 10 can be interpreted as the effect of a 10% difference in BW on % SD-outcomes; controls: birth order for all models and grade track for z-academic performance; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10

It should be noted, though, that the differences in the BW coefficients by SES reported in Table 4, as shown in the columns labelled as "Low–High" testing for a joint interaction, are not statistically significant. The coefficients shown in the "Low-High" column of Table 4 express the difference in the coefficients of BW between low-educated families (Low-Edu. column) and high-educated families (High-Edu. column), which are estimated in independent models. The reference category is the effect of BW on outcomes for high-educated families. The SES-differences are only significant at 10% with a one-tailed *t-test* assuming that, according to theoretical predictions and previous findings, the effect should be larger for low-SES families. This raises some concerns about a lack of statistical power due to small sample sizes and attenuation bias in twin-fixed effects models (McGue et al. 2010).

To put these effects of BW in perspective, for instance, its effect (10% SD units) on academic performance among low-educated families for the 11-year-old cohort accounts for 21% of the SES-gap in academic performance (0.48 SD unit difference between high- and low-educated families; see Table A.1. and Figure A.1.). Previous research found that a 10% twin difference in BW is associated with an increase by about 5% SD units in school grades, whilst large-scale educational interventions in developing countries report increases in test scores between 0.17 and 0.47 SD (Bharadwaj et al. 2018:351).

6.1. Mediation: Within-Family Allocation of Investments

To test whether and how parents allocate investments within-families explains the observed patterns by SES: Table 5 displays the results of the mediation analysis by academic outcome, parenting measure and SES. We find that, on average, low-SES families do not *respond* to their children's birth endowments but allocate their investments evenly irrespectively of BW. Parenting coefficients are neither statistically significant nor substantial in magnitude, ranging from 0.5 for cultural activities and 0.3 for warmth. Given that there is no substantial link between BW and parental investments, the latter cannot play a mediating role in reinforcing the effect of BW differences on academic outcomes.

Table 5. Twin-fixed effects (FE) OLS models for the effect of log(BW) on parental investments by parental education for the 11-year-old cohort

Investment Outcomes	Cultural Activities				Parental Warmth			
	All Sample	**Low-Edu.**	**High-Edu.**	**Low-High**	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-High**
Log(birth weight)	0.555**	0.469	0.680*	-0.179	0.0732	0.331	-0.280	0.614
	(0.267)	(0.382)	(0.354)	(0.514)	(0.206)	(0.273)	(0.303)	(0.404)
Observations	1,530	812	718	1,530	1,578	806	772	1,578

Notes: controls: birth order; robust standard errors between parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10

By contrast, high-SES families tend to spend more time in cultural activities with the heavier-BW co-twins, so slightly reinforcing their allocation of investments. A 10% difference in BW is related to about 7% SD units more time spent in cultural activities with the heavier-BW co-twins (*p-value* at 0.10%). For parental warmth, though, we do not find parental response to twins' differences in BW. Consequently, parental allocation of investments cannot play a mediating role in compensating for BW differences among high-SES families.

We argue that these results on inconsistent mediation can be interpreted as a preliminary indication that the within-family allocation of resources or investments does not account for the observed patterns by SES—no effect of BW for high-SES families and effect for low-SES families– that we saw above in Table 4. The general finding on neutral *parental response* in terms of cultural activities and warmth to twins' BW differences is in line with previous findings from Chilean twins (Bharadwaj et al. 2018; Abufhele et al. 2017).

6.2. Moderation: Absolute Investments Between Families

Table 6 shows descriptive evidence on the relative allocation of investments within families and the absolute level of investments according to the lighter- or heavier-BW twin in each family. As can be seen, at 5 years old, both low and high-SES families invest equally irrespectively of the BW of their children. At 11 years old, low-SES families still invest equally, while high-SES families invest slightly more in the heavier-BW co-twins (reinforcing parental response). This descriptive evidence is in line with the previous mediation analysis showing that high-SES families are slightly more prone to reinforce their investments.

Table 6. Parental involvement in cultural activities (upper panel) and parental warmth (lower panel) by parental education, twins' BW and birth cohort

	Mean Parental Involvement in Cultural Activities (z-scores)					
	Cohort: 5-year-old			**Cohort: 11-year-old**		
Within Family Birth Weight	**Low-Edu. (1)**	**High-Edu. (2)**	**(1-2)**	**Low-Edu. (1)**	**High-Edu. (2)**	**(1-2)**
Lower Weight co-twin	-0.17	0.19	-0.35***	-0.14	0.07	-0.21***
Heavier Weight co-twin	-0.14	0.20	-0.34***	-0.09	0.22	-0.31***
Total	-0.16	0.20	-0.36***	-0.11	0.14	-0.25***
n	996	804		812	718	

	Mean Parental Warmth (z-scores)					
	Cohort: 5-year-old			**Cohort: 11-year-old**		
Within Family Birth Weight	**Low-Edu. (1)**	**High-Edu. (2)**	**(1-2)**	**Low-Edu. (1)**	**High Edu. (2)**	**(1-2)**
Lower Weight co-twin	-0.07	0.01	-0.03*	-0.11	0.13	-0.25***
Heavier Weight co-twin	0.01	0.03	-0.02	-0.10	0.13	-0.22***
Total	-0.03	0.02	-0.06	-0.11	0.13	-0.24***
n	1,008	812		806	772	

Notes: Two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10

The most interesting pattern coming out of Table 6 is that, in the 11-year-old cohort, high-SES families invest more in absolute terms in both the lighter-BW and heavier-BW co-twins than low-SES families. For instance, high-SES families spend 0.07 SD units in cultural activities with lower-BW co-twins (-0.15 SD less than in heavier-BW ones), while low-SES families -0.14 SD units, a 1/5 SD difference. Overall, in absolute terms, high-SES families invest ¼ SD units more in cultural activities and warmth than low-SES families.

Table 7. Twin-FE OLS models of the effect of log(BW) on educational outcomes by parental education moderated by between-family average quintiles of parental investments for the 11-year-old cohort

Outcome	**Academic Performance**		**Cognitive Ability**		**Behavioural Problems**	
Investment	**Cultural Activities**	**Parental Warmth**	**Cultural Activities**	**Parental Warmth**	**Cultural Activities**	**Parental Warmth**
			All Sample			
Log(birth weight)	1.802***	0.671	-0.0613	1.238**	-0.0788	-1.782***
	(0.563)	(0.494)	(0.573)	(0.574)	(0.535)	(0.685)
Log(birth weight) X q2 investment	-1.354*	0.339	0.0763	-0.655	-0.820	1.093
	(0.704)	(0.651)	(0.792)	(0.744)	(0.762)	(0.795)
Log(birth weight) X q3 investment	-1.982**	-0.604	-0.845	-1.409*	0.283	2.228**
	(0.855)	(0.686)	(0.867)	(0.805)	(0.845)	(0.885)
Log(birth weight) X q4 investment	-1.647**	0.133	0.622	-1.606**	-0.389	1.765**
	(0.644)	(0.649)	(0.707)	(0.728)	(0.683)	(0.854)
Log(birth weight) X q5 investment	-0.309	-0.0441	0.620	-1.190	-0.580	1.672*
	(0.866)	(0.876)	(0.848)	(0.765)	(0.831)	(0.925)
Observations	1,285	1,262	1,532	1,521	1,659	1,628

Notes: controls: birth order for all models and grade track for z-academic performance; robust standard errors between parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10

To formally test the moderating (neutralising) role of the absolute level of parental investments (twin-pair mean), in Table 7 we estimate an interaction between BW and two measures of parental investment or involvement (cultural activities and warmth). The reference category is the first row labelled "Log(BW)", corresponding to the (fixed) effect of BW on academic outcomes among families at the lowest absolute level of parental investments (bottom quintile). As can be seen, overall, there are statistically significant interaction effects by which the effect of BW is largest at the bottom quintile of parental investments. For academic performance, we find a moderating role of parental time in cultural activities with their children, while for cognitive ability and behavioural problems, parental warmth moderates the effect of BW.

Even when there are nonlinear effects at certain quintiles that suggest caution with our interpretation, the general pattern is that, from the second quintile of absolute parental investments, the effect of BW is reduced or even neutralised, while at the lowest quintile of parental investments we tend to find the largest effect-size of BW. However, we cannot disentangle why parental involvement in cultural activities and warmth have a different moderating role depending on the academic outcome analysed.

According to the proposed mechanism on *absolute resources/investments between families*, even when high-SES families tend to invest slightly more in heavier-BW children, their large absolute level of resources allows them to deploy high-quality investments that tend to neutralise the effects of health shocks on later skill formation. In contrast, as low-SES families invest a lower amount of absolute investments in both lighter-BW and heavier-BW children and are allocated in a steeper section of the skill function (see Figure 2 above), twin differences in BW predict later differences in academic skills even when, as we saw above, within-family resource allocation is neutral.

As highlighted by the literature from developmental psychology and the findings from educational interventions, positive parent-child interactions, and cognitive stimulation are important factors for children's cognitive and behavioural development. Consequently, we argue that biology is not destiny because (enriched) social environments might offset the detrimental effect of prenatal health shocks, such as LBW, on early skill formation.

6.3. Robustness Checks

Throughout the article, we carried out several robustness checks[39] to assess the credibility of the findings on the following: (1) alternative specifications of parental SES; (2) alternative specifications of birth endowments (foetal growth and prematurity); (3) heterogeneity by twin-pair zygosity and gender; (4) alternative Naïve OLS models; (5) nonlinearities across the absolute BW distribution and prenatal confounding; and (6) further mechanisms on health problems and schooling investments. In the remainder of this chapter, we comment extensively on the fifth and sixth robustness checks.

6.3.1. Nonlinearities Across Birth Weight Distribution and Prenatal Confounding

Even though twin-FE models control for most sources of unobserved confounding to identify the causal effect of BW, when assessing SES-heterogeneity one could argue that within-family differences in BW are drawn from absolute BW distributions that differ by SES. That is to say that there may be issues of endogeneity with the measure of BW, as it might also proxy for prenatal investments in addition to birth endowments.

As shown in Figure 3 (left-hand side) above, low-SES families have a larger prevalence of LBW (≈6%) that may capture SES-differences in prenatal behaviour and access/quality of

[39] Some analyses and robustness checks are only shown for the 11-year-old cohort because I found a heterogeneous effect of BW by parental SES for this cohort only.

antenatal care. Thus, within-family differences in BW may be more detrimental for low-SES than high-SES families not only due to SES-gaps in postnatal investments but also because the former have a larger prevalence of LBW, which is negative for child development. To account for this possibility, in Appendix Table A.5., we carry out a robustness check across the absolute BW distribution and parental SES among families under or above the LBW threshold. Additionally, we assess prenatal parental behaviour, health conditions and antenatal care by SES in Tables A.11–A.12.

As can be seen in Table A.12., low-SES families are around 6% more likely to deliver LBW twins (and preterm twins at ≈3%) than high-SES families, a difference related to SES-gaps in prenatal investments such as visiting the hospital just before giving birth. However, we did not find further SES-differences during the week of the initial examination or the number of preventive examinations during pregnancy. Moreover, in Table A.11. we analyse the medical record of most common risk factors for the pregnancy by SES, and only non-remarkable differences were found. Unfortunately, information on maternal smoking and nutrition was not available.

Differences by SES in prenatal factors associated with LBW—e.g., antenatal access/quality—do not rule out the possibility that (lack of) postnatal investments complement (lack of) prenatal investments by reinforcing the effect of BW, as we tend to find negative effects of BW only among low-SES and LBW children.[40] As the negative effect of LBW for skill formation is a well-established finding in the epidemiological literature, its smaller/null effect among high-SES families suggests that SES heterogeneity in the effect of BW is not only driven by prenatal factors or it is spurious.

By the same token, in the moderation analysis by family-level investments shown in *section 6.2.*, between-family confounding may be an issue if families at different investment quintiles differ systematically in prenatal characteristics other than SES: as expected, families at the bottom quintile of investments are lower-educated and have a lower ISEI. We explored the characteristics of the subsamples by quintiles of investments in Table A.15. As can be seen, families at the bottom quintile of investments, where the effect-size of BW is largest, are more likely to deliver LBW twins and mothers are younger with respect to families at second-to-fifth quintiles, but there are no other considerable differences in terms of prematurity, or

[40] We also apply the same heterogeneity analysis by SES and prematurity, and results are similar to LBW (see Appendix Table A.8.).

antenatal care. Still, as parental investments are not randomly allocated and are measured after the health shock, we cannot make a causal claim. Future research designs may shed further light on the causal role of environmental factors in moderating health shocks by exploiting natural experiments inducing exogenous variation in both factors—*a lightning strikes twice* (Almond et al. 2018).

6.3.2. Additional Mechanisms: Health Problems

In order to explore biological mechanisms underlying the observed causal effect of BW on skill formation, we explore the role of health problems or diagnosis of physical/mental illnesses/disability as a plausible mediator. About 56% of twins were diagnosed with at least one physical or mental illness[41] up to age 11, with no variation by parental SES. In Table A.9. we run twin-FE linear probability models to assess the effect of BW on health problems. At age 11, a 10% difference in BW is associated with a 2% (*p-value* at 10%) increase in the likelihood of being diagnosed with at least one illness. However, consistently with the main results from Table 4, we only find an effect of BW on health problems among low-educated families (β = -0.28; *p-value* at 5%) and a null effect among high-educated families (β = -0.05; *p-value*>10%) —this holds even when estimating models among LBW twins. This finding suggests that BW may be more detrimental for skill formation among low-SES families due to children's health problems likely related to parental resources and investments (e.g., health care quality).

6.3.3. Additional Mechanisms: School Investments

School quality, access to compensatory education, and within-family allocation of schooling investments are complementary mechanisms to parenting when it comes to moderate the effect of health shocks across family SES (Aizer and Cunha 2012). Enriching schooling environments with tailored support for kids with special educational needs and disabilities, such as LBW twins, may be particularly effective to neutralise adverse early-life conditions (Baranowska-Rataj et al. 2019). To shed some light on the schooling mechanism, we assessed whether high- and low-SES families have different levels of access to preschool[42] and school[43]

[41] Respiratory illness, allergies, neurodermatitis, defective vision, eating disorder, motor dysfunction, mental disability, physical disability, anxiety disorder, social behaviour disorder, attention deficit disorder, dyslexia (reading/writing difficulties), dyscalculia (difficulties with mathematics), stuttering, other physical or mental illness.

[42] Twin-specific use (ICC=0.72) of the support provided by kindergarten or another institution for special needs education in one or more of the following programs: learning, speaking, physical and motor development, emotional and social development, mental development, vision, hearing, autism and/or others.

[43] Twin-specific participation (ICC=0.85) outside regular school hours in one or more of the following activities: help with homework; remedial groups; and/or subject-specific additional classes.

compensatory programs, and general preschool activities[44]. As shown in Table A.13., we find that low-SES families are more likely to get access to compensatory educational programs in both preschool and school, even after controlling for LBW. Regarding average attendance in preschool activities, which can be well-considered as a proxy for preschool quality, highly-educated families have an advantage of ¼ SD in comparison to low-educated families. Thus, high-SES families tend to enrol their children in pre-school institutions that offer academic activities more frequently, which could add up to their observed high levels of resources and parental involvement. Unfortunately, no direct indicators of preschool or school quality are available in the *Twin Life* dataset.

Finally, we analysed whether parents allocate schooling investments differently among twins as a function of BW. As displayed in Table A.14. left-hand panel, both low- (β = -0.28; *p-value* at 5%) and high-educated (β = -0.22; *p-value>10%*) families tend to compensate the disadvantage of lighter BW co-twins by enrolling them more into compensatory preschool education. However, as shown in Table A.14. right-hand panel, high-SES families tend to reinforce schooling investments by enrolling the heavier-BW co-twins more into compensatory programs during elementary education (β = 0.24; *p-value* at 5%), while no differences are found among low-SES families. This pattern is in line with the previous analysis on within-family differences in cultural activities in which high-SES families are slightly more prone to reinforce.

The inconsistent evidence on null SES differences in access to compensatory programs, and high-SES reinforcement of schooling investments, suggest that SES heterogeneity in the effect of BW is not likely explained by unequal schooling access and/or allocation of schooling investments. Better measures of preschool and school quality may help to test if school quality may actually complement the role of parenting across families.

7. Conclusions

The main aim of this article was testing the effect of BW on skill formation at age 5 and 11 by bridging the literature on social stratification, human capital formation, developmental psychology and epidemiology. Birth weight is a good indicator of birth endowments and a

[44] Child taking part (ICC=0.98) during the entire time in childcare up to the age of 6 (normal kindergarten and Kita activities as well as special activities beyond the standard program) in the following activities: early musical education, drawing/painting, help with learning German, foreign language classes, mathematical and scientific stimuli, visits to libraries and nature trips.

predictor of children's early development and socioeconomic attainment later in life. However, the consequences of BW are not biological destiny, since enriched social environments may neutralise its effect. This potential socioeconomic gradient in the effect of BW on skill formation might account for the persistence of early SES-gaps in the process of human capital accumulation.

The main contribution of this article lied in assessing the stratification of the association between BW and skill formation by parental SES and scrutinising its potential mechanisms. We predicted that high-SES families are more likely to compensate for the negative effect of prenatal health shocks than low-SES families, given their extensive pool of cultural and economic resources. To test this *compensatory hypothesis*, we used a twin-FE design that allowed us to identify the causal effect of BW by exploiting random sources of variation within-families (e.g., intrauterine foetal growth).

The article was mainly motivated by mixed results from previous research and the limited number of studies that have explicitly focused on the heterogeneity of health shocks by parental SES. We further contributed to the literature by exploring two possible mechanisms that may account for the heterogeneous effect of BW by SES: (1) relative allocation of investments within families; and (2) absolute level of resources between families. In accordance with the literature on developmental psychology and sociology, we measured the level of investments of the families with two indicators of positive parenting related to children's well-being and academic performance: parental time in cultural activities and emotional warmth.

Our second contribution lied in exploring the effect of BW on three key outcomes for academic success: cognitive ability, behavioural problems, and academic performance in mathematics and language. Thus, we contributed to the literature on the multidimensionality of human capital formation.

Results showed that lower-BW co-twins have worse academic performance and more behavioural problems than their heavier-BW co-twins. At age 5, we observe a causal effect of BW on academic performance and behavioural problems for high- and low-SES families alike. This effect of BW fades away (or it is reduced) for children of high-SES parents by age 11. We argue that this pattern of *compensatory advantage* among high-SES families in the 11-year-old cohort may be explained by their high absolute level of resources and investments, but not by its relative allocation of investments within-families. Therefore, we argue that biology is not destiny because (enriched) social environments may offset the detrimental effect of BW on early skill formation.

Even when we cannot directly compare results across our birth cohorts aged 5 and 11, the observed pattern of compensation or null effect of BW among high-SES families at age 11—in comparison to its negative effect at age 5—is puzzling. Theories of human capital formation predict that interventions or environmental input are more productive in neutralising health shocks during sensitive or critical stages of early child development. Alternatively, it could be the case that, as variation in the complexity of skill formation increases from early childhood, high-SES parents may have more room to compensate for the detrimental effect of BW across pre-school and elementary education.

We carried out several robustness checks to assess the credibility of the findings by using alternative specifications of parental SES and BW, exploring nonlinearities in the effect of BW across its absolute distribution, testing for prenatal and genetic confounding, and exploring further biological and environmental mechanisms such as health problems and schooling investments, to generally conclude that the study's main findings are consistent.

Still, the study has four limitations that should be addressed in future research. First, as most twin studies do, we deal with low sample sizes that limit statistical power to find statistically significant differences by subgroups. Moreover, twin-FE models suffer from attenuation bias due to measurement error (Kohler et al. 2011). One way of overcoming these limitations is to use administrative data with very large sample sizes and more reliable measures.

Second, external validity in twin studies is a general issue. We wonder to what extent twin differences in BW that are concentrated at the bottom of the BW distribution (around 50% of twins are LBW) are informative for the whole population of single births. The main source of variation in BW (intrauterine foetal growth) among twins only accounts for around 30% of the incidence of LBW in the population of singletons, which mainly comes from gestational age (Baranowska-Rataj et al. 2019).

Third, even though we find considerable within-family variation in parental allocation of investments, it may be particularly difficult for parents to differentiate (and report) their investments among twins due to inequity aversion and *spillover* and *common goods effects.* Datasets combining twins and siblings would be very helpful in assessing whether this is actually the case. Furthermore, the measures of parental investments are observed at an unspecified retrospective window before the survey, which is particularly problematic for the school grades outcome that also refer to the past. Thus, endogeneity could be an issue when assessing the moderating effect of parental investments on academic outcomes, as the former might be responsive to previous skills.

Fourth, we cannot interpret our findings—causal effect of BW in the 5-year-old cohort for both low and high-SES families, and null effect in the 11-year old cohort for high-SES families—longitudinally as we analysed two different birth cohorts. Only under the untestable but plausible assumptions that (1) these birth cohorts were comparable in all their observed and unobserved characteristics (see Appendix Table A.1.), and that (2) there were no cohort or period effects, could we do so. Thus, future studies drawing from longitudinal twin data would shed further light on the effect of prenatal health shocks on skill formation across different stages of child development.

This article has provided substantive contributions in theoretical, methodological and empirical terms to an emerging interdisciplinary literature on socioeconomic inequalities on the consequences of perinatal health for early skill formation while acknowledging the limitations discussed above as areas for improvement in future research. The parenting mechanisms that we identify as positive for child development in offsetting prenatal health shocks may inform future educational interventions targeted at LBW children. The scarce available evidence shows that the most-effective interventions were based on intense psycho-social stimulation by parents and trained professionals (weekly home visits or centre-based) among LBW infants during the first three years of life. However, this field of research is in its infancy, and further evidence is needed to disentangle the complex biological, socioeconomic, and behavioural determinants of LBW and to determine how to mitigate their negative consequences for child development.

APPENDIX

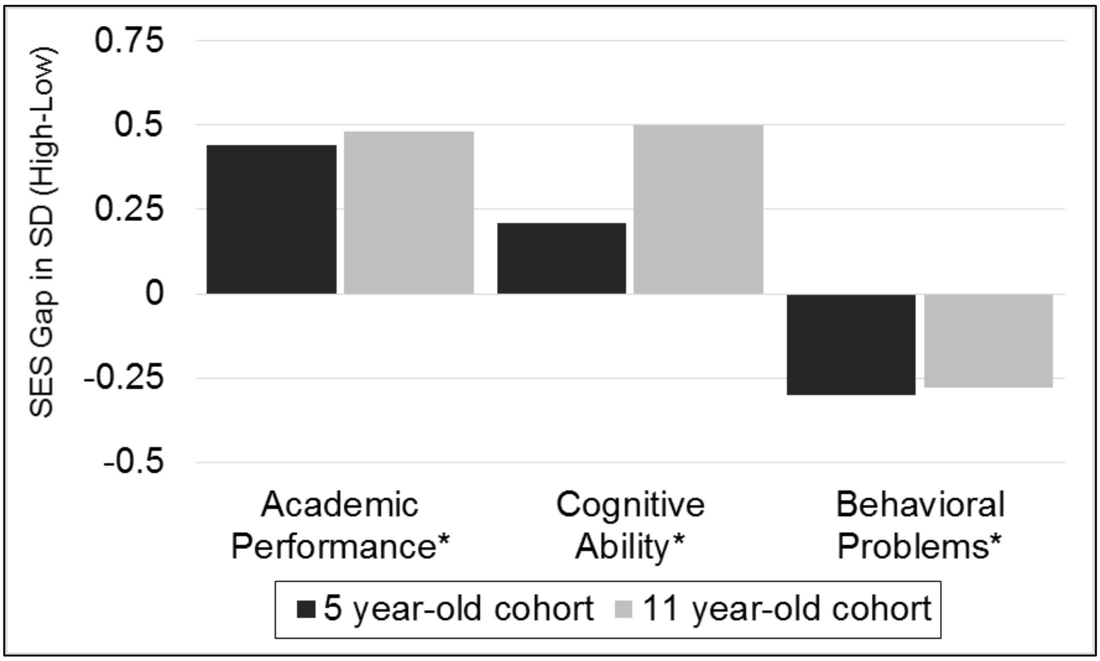


Figure A.1. SES gap in academic skills by birth cohort
Notes: *SES-differences statistically significant at p-value<0.01 (two-tailed t-test)

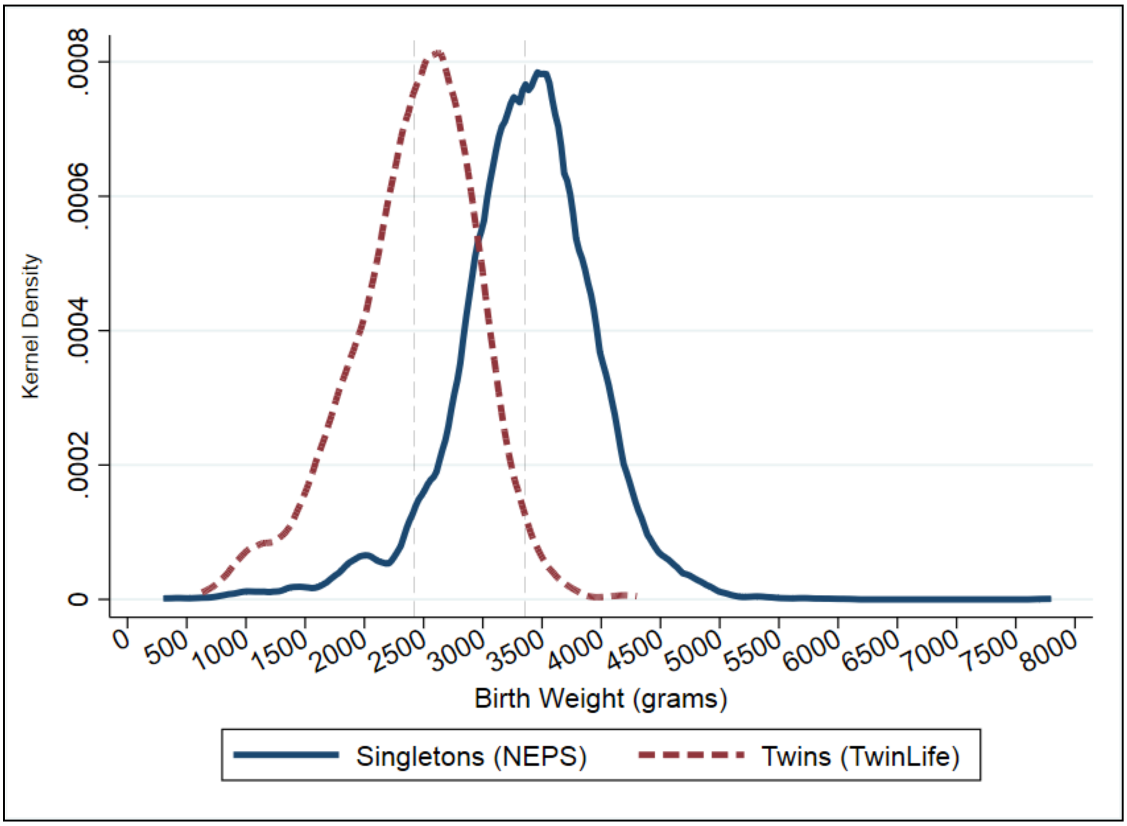


Figure A.2. Distribution of BW in twin and singleton samples

Table A.1. Descriptive statistics by birth cohort and parental education

5-year-old Cohort												
SES	**Low Education**						**High Education**					
Variables	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**ICC**	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**ICC**
Female twin-pair	1,118	0.51		0	1		880	0.52		0	1	
Dizygotic twin-pair	1,118	0.56		0	1		880	0.58		0	1	
Birth order	1,118	1.5		1	2		880	1.5		1	2	
Age in months	1,118	65.48	2.95	52	75		880	65.38	3.77	57	74	
Mother's age	1,112	36.08	5.43	22	59		878	38.83	4.20	26	50	
Highest parental ISEI	1,030	37.78	22.84	0	88		878	73.17	19.51	0	89	
Parental involvement in cultural activities	1,086	2.34	0.40	1.03	3.26	0.96	862	2.48	0.33	1.20	3.26	0.98
z - Parental involvement in cultural activities	1,086	-0.16	1.07	-3.67	2.31	0.96	862	0.20	0.86	-3.14	2.31	0.98
z - Within-family differences in cultural activities	1,086	0.20	0.35	0	3.27		862	0.12	0.22	0	1.76	
Parental warmth (raw)	1,102	4.55	0.37	3	5	0.89	872	4.56	0.37	3	5	0.94
z - Parental warmth	1,102	-0.01	0.99	-4.16	1.19	0.89	872	0.02	1.00	-4.16	1.19	0.94
z - Within-family differences in parental warmth	1,102	0.35	0.50	0	5.35		872	0.29	0.40	0	3.35	
Birth weight (grams)	1,020	2,320.87	565.97	280	3,880	0.88	820	2,396.48	545.03	475	3,520	0.89
Within-family birth weight differences (grams)	1,020	275.35	240.45	0	1,920		820	268.92	218.16	0	1,050	
Log(birth weight)	1,020	7.71	0.29	5.63	8.26		820	7.75	0.27	6.16	8.17	
LBW	1,020	0.60		0	1	0.70	820	0.53		0	1	0.69
z – Foetal growth	1,020	-0.05	1	-5.20	2.95	0.60	820	0.07	0.97	-3.17	2.81	0.58
Gestation week	1,020	35.52	2.72	25	44		820	35.72	2.69	22	42	
Preterm	1,020	0.57		0	1		820	0.53		0	1	
Physical/mental illness diagnosis	1,102	0.47		0	1	0.74	854	0.48		0	1	0.61
z -Cognitive ability	830	-0.04	0.99	-2.32	3.43	0.78	682	0.17	1.00	-2.46	3.53	0.78
z -Academic performance	1,096	-0.19	1.00	-3.14	2.34	0.74	858	0.25	0.94	-3.14	2.34	0.67
Behavioural problems (raw scores)	1,092	8.14	4.87	0	30	0.66	850	6.71	4.16	0	22	0.57
z -Behavioural problems	1,092	0.12	1.06	-1.47	5.90	0.69	850	-0.18	0.87	-1.47	3.13	0.57

Table A.1. Continued

11-year-old Cohort												
SES	**Low Education**						**High Education**					
Variables	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**ICC**	**Obs.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**ICC**
Female twin-pair	1,108	0.55		0	1		962	0.49		0	1	
Dizygotic twin-pair	1,108	0.57		0	1		962	0.63		0	1	
Birth order	1,108	1.5		1	2		962	1.5		1	2	
Age in months	1,108	137.43	3.74	130	146		962	137.63	3.64	130	147	
Mother's age	1,100	41.75	4.94	28	54		960	44.45	4.39	30	58	
Highest parental ISEI	994	36.35	20.56	0	81		950	70.80	19.80	0	89	
Parental involvement in cultural activities	898	1.46	0.52	0.66	3.30	0.54	780	1.60	0.52	0.66	3.30	0.61
z - Parental involvement in cultural activities	898	-0.11	1.00	-1.64	3.64	0.54	780	0.14	0.98	-1.64	3.64	0.60
z - Within-family differences in cultural activities	898	0.83	0.76	0	4.41		780	0.80	0.67	0	3.73	
Parental warmth (raw)	884	4.29	0.66	1.5	5	0.65	838	4.45	0.53	2.38	5	0.57
z - Parental warmth	884	-0.11	1.05	-4.56	1.03	0.65	838	0.16	0.85	-3.16	1.03	0.57
z - Within-family differences in parental warmth	884	0.78	0.76	0	4.59		838	0.71	0.62	0	3.00	
Birth weight (grams)	1,004	2,378.32	550.37	630	4,220	0.88	882	2,467.65	524.46	720	4,300	0.86
Within-family birth weight differences (grams)	1,004	266.72	248.92	0	1,440		882	272.60	253.00	0	1,800	
Log(birth weight)	1,004	7.74	0.27	6.45	8.35		882	7.78	0.24	6.58	8.37	
LBW	1,004	0.55		0	1	0.68	882	0.47		0	1	0.74
z – Foetal growth	1,004	-0.05	1.01	-3.84	4.26	0.61	882	0.07	0.96	-2.75	4.91	0.61
Gestation week	1,004	35.71	2.63	24	42		882	35.99	2.46	24	41	
Preterm	1,004	0.55		0	1		882	0.51		0	1	
Physical/mental illness diagnosis	1,090	0.55		0	1	0.63	936	0.57		0	1	0.64
z -Cognitive ability	1,000	-0.21	1.01	-3.36	2.86	0.66	884	0.29	0.90	-2.82	2.60	0.60
z -Academic performance	794	-0.21	1.01	-3.23	2.00	0.74	744	0.27	0.90	-2.58	2.00	0.73
Behavioural problems (raw scores)	1,080	10.79	5.44	0	30	0.52	922	9.45	4.77	0	27	0.46
z -Behavioural problems	1,080	0.13	1.06	-1.78	4.12	0.53	922	-0.15	0.91	-1.78	3.90	0.49

Table A.2. Naïve LPM on missing data by outcome and parenting measure for the 11-year-old cohort

Missing Data	Academic Performance			Cognitive Ability			Behavioural Problems		
	Overall	Cultural	Warmth	Overall	Cultural	Warmth	Overall	Cultural	Warmth
High parental education	-0.0656**	-0.0515*	-0.112***	-0.0241	-0.0146	-0.0711***	0.0223*	0.0116	-0.0641***
	(0.0282)	(0.0284)	(0.0293)	(0.0190)	(0.0225)	(0.0242)	(0.0122)	(0.0177)	(0.0205)
Female twin-pair	-0.0196	-0.0260	-0.0171	-0.0348*	-0.0469**	-0.0185	0.00703	-0.0190	0.0183
	(0.0283)	(0.0284)	(0.0294)	(0.0186)	(0.0222)	(0.0242)	(0.0117)	(0.0175)	(0.0207)
Dizygotic twin-pair	0.0403	0.0228	0.0396	-0.0105	-0.00479	-0.00593	0.00004	0.00414	0.00814
	(0.0286)	(0.0289)	(0.0298)	(0.0195)	(0.0232)	(0.0249)	(0.0120)	(0.0180)	(0.0211)
Log(birth weight 1st born)	0.0191	-0.0535	0.0865	0.00437	-0.0870	0.0636	-0.0257	-0.106	0.0319
	(0.102)	(0.0996)	(0.101)	(0.0582)	(0.0769)	(0.0802)	(0.0379)	(0.0654)	(0.0702)
Log(birth weight 2nd born)	-0.152	-0.105	-0.230**	-0.0321	-0.0353	-0.140*	0.00718	0.00302	-0.106
	(0.100)	(0.0976)	(0.0982)	(0.0565)	(0.0739)	(0.0770)	(0.0304)	(0.0603)	(0.0652)
Constant	1.296***	1.575**	1.482***	0.340	1.171***	0.831**	0.162	0.924***	0.728**
	(0.492)	(0.474)	(0.486)	(0.309)	(0.386)	(0.403)	(0.165)	(0.315)	(0.331)
n	1,886	1,886	1,886	1,886	1,886	1,886	1,886	1,886	1,886
n missing	474	601	624	170	353	365	60	227	258
R-squared	0.015	0.011	0.025	0.006	0.010	0.013	0.005	0.007	0.014

Robust Standard errors in parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10; *n=1,886* is the baseline analytic sample after applying list-wise deletion (*n missing=196*) to the original sample for parental education, gender, zygosity and BW.

Table A.3. Twin-FE OLS models for the effect of log(BW) on educational outcomes by birth cohort and parental ISEI

	5-year-old cohort				11-year-old cohort			
Academic Outcomes	**All Sample**	**Low-ISEI**	**High-ISEI**	**Low - High**	**All Sample**	**Low-ISEI**	**High-ISEI**	**Low - High**
z - Academic performance	0.64***	0.55*	0.75**	-0.19	0.69***	0.99***	0.48*	0.46
	(0.22)	(0.28)	(0.32)	(0.43)	(0.23)	(0.32)	(0.29)	(0.45)
n	1,724	850	800	1,724	1,324	676	648	1,324
z - Cognitive abilities	0.05	-0.25	0.41	-0.63*	0.16	0.42	-0.18	0.58
	(0.16)	(0.24)	(0.26)	(0.35)	(0.23)	(0.30)	(0.34)	(0.45)
n	1,340	644	696	1,340	1,610	840	770	1,610
z - Behavioural problems	-0.68***	-0.44*	-0.97***	0.55	-0.30	-0.59*	-0.08	-0.69
	(0.19)	(0.26)	(0.26)	(0.40)	(0.24)	(0.34)	(0.33)	(0.47)
n	1,714	846	868	1,714	1,718	914	804	1,718

Notes: FE=OLS fixed effects; the coefficients multiplied by 10 can be interpreted as the effect of a 10% difference in BW on %SD of the outcome; controls: birth order for all models and grade track for z-academic performance; robust standard errors between parentheses;
two-tailed t-test: ***p-value<0.01; **p-value<0.05; * p-value<0.10

Table A.4. Twin-FE OLS models for the effect of log(BW) on parental investments by parental ISEI for the 11-year-old cohort

Investment Outcomes	Cultural Activities				Parental Warmth			
	All Sample	Low-ISEI	High-ISEI	Low - High	All Sample	Low-ISEI	High-ISEI	Low - High
Log(birth weight)	0.542**	0.556	0.519	0.013	0.0992	0.310	-0.176	0.449
	(0.268)	(0.382)	(0.363)	(0.517)	(0.214)	(0.281)	(0.327)	(0.428)
Observations	1,440	762	678	1,440	1,486	760	726	1,486

Notes: controls: birth order; robust standard errors between parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10

Table A.5. Twin FE-models by birth cohort, outcome, parental education and low (LBW) or normal birth weight (NBW)

5-year-old cohort						
Academic Performance						
	All Sample	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-Edu.**	**High-Edu.**
	LBW (<2,500gr)	**NBW (>=2,500gr)**	**LBW (<2,500gr)**	**LBW (<2,500gr)**	**NBW (>=2,500gr)**	**NBW (>=2,500gr)**
Log(birth weight)	0.150	1.245***	-0.0541	1.400***	0.708*	0.973**
	(0.239)	(0.272)	(0.270)	(0.322)	(0.418)	(0.483)
n	792	959	421	560	371	399
Cognitive Ability						
	All Sample	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-Edu.**	**High-Edu.**
	LBW (<2,500gr)	**NBW (>=2,500gr)**	**LBW (<2,500gr)**	**LBW (<2,500gr)**	**NBW (>=2,500gr)**	**NBW (>=2,500gr)**
Log(birth weight)	-0.00728	0.203	-0.122	-0.155	0.291	0.699
	(0.202)	(0.273)	(0.239)	(0.348)	(0.493)	(0.439)
n	631	729	330	415	301	314
Behavioural Problems						
	All Sample	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-Edu.**	**High-Edu.**
	LBW (<2,500gr)	**NBW (>=2,500gr)**	**LBW (<2,500gr)**	**LBW (<2,500gr)**	**NBW (>=2,500gr)**	**NBW (>=2,500gr)**
Log(birth weight)	-0.735***	-0.703**	-0.587***	-0.664*	-1.170**	-0.764*
	(0.210)	(0.295)	(0.224)	(0.389)	(0.530)	(0.456)
n	792	949	421	556	371	393

Notes: robust standard errors in parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10;
Controls: birth order for all models.

Table A.5. Continued

11-year-old cohort						
Academic Performance						
	All Sample	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-Edu.**	**High-Edu.**
	LBW (<2,500gr)	**NBW (>=2,500gr)**	**LBW (<2,500gr)**	**LBW (<2,500gr)**	**NBW (>=2,500gr)**	**NBW (>=2,500gr)**
Log(birth weight)	1.037***	0.137	1.377***	0.708**	0.221	0.0333
	(0.255)	(0.397)	(0.408)	(0.290)	(0.551)	(0.542)
n	710	702	392	318	328	374
Cognitive Ability						
	All Sample	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-Edu.**	**High-Edu.**
	LBW (<2,500gr)	**NBW (>=2,500gr)**	**LBW (<2,500gr)**	**LBW (<2,500gr)**	**NBW (>=2,500gr)**	**NBW (>=2,500gr)**
Log(birth weight)	0.371	-0.259	0.562*	0.00954	-0.199	-0.343
	(0.267)	(0.358)	(0.339)	(0.428)	(0.567)	(0.463)
n	856	860	492	364	410	450
Behavioural Problems						
	All Sample	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-Edu.**	**High-Edu.**
	LBW (<2,500gr)	**NBW (>=2,500gr)**	**LBW (<2,500gr)**	**LBW (<2,500gr)**	**NBW (>=2,500gr)**	**NBW (>=2,500gr)**
Log(birth weight)	-0.321	-0.456	-0.628*	0.300	-0.583	-0.360
	(0.273)	(0.436)	(0.337)	(0.476)	(0.761)	(0.500)
n	904	922	532	372	450	472

Notes: robust standard errors in parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10;
Controls: birth order for all models and grade track for z-academic performance.

Table A.6. Twin FE-models by birth cohort, outcomes, twin-pair zygosity and gender

		5-year-old cohort						11-year-old cohort					
		Academic Performance		Cognitive Ability		Behavioural Problems		Academic Performance		Cognitive Ability		Behavioural Problems	
		DZ-Twins	**MZ-Twins**	**DZ-Twins**	**MZ-Twins**	**DZ-Twins**	**MZ-Twins**	**DZ-Twins**	**MZ-Twins**	**DZ-Twins**	**MZ-Twins**	**DZ-Twins**	**MZ-Twins**
Log(birth weight)		0.622**	0.741***	-0.00544	0.230	-0.758***	-0.459	0.701**	0.748***	-0.0396	0.528*	-0.139	-0.788**
		(0.290)	(0.226)	(0.168)	(0.357)	(0.211)	(0.386)	(0.324)	(0.245)	(0.277)	(0.331)	(0.304)	(0.327)
	n	1,024	778	800	596	1,018	774	828	584	1,028	688	1,088	738
		Males	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
Log(birth weight)		1.199***	0.300	0.323	-0.115	-0.636*	-0.668***	0.687**	0.725**	-0.318	0.565**	0.0556	-0.752**
		(0.397)	(0.210)	(0.296)	(0.196)	(0.326)	(0.229)	(0.289)	(0.326)	(0.325)	(0.285)	(0.349)	(0.301)
	n	878	924	670	726	876	916	680	732	820	896	892	934

Notes: robust standard errors in parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10; Controls: birth order for all models and grade track for z-academic performance.

Table A.7. Twin-FE OLS models for the effect of z - foetal growth on educational outcomes by birth cohort and parental education

	5-year-old cohort				11-year-old cohort			
Academic Outcomes	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low - High**	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low-High**
z - Academic performance	0.115***	0.131***	0.0951**	0.041	0.116***	0.166***	0.0750*	0.086
	(0.0314)	(0.0429)	(0.0457)	(0.0619)	(0.0349)	(0.0523)	(0.0445)	(0.068)
n	1,799	999	800	1,799	1,412	720	692	1,412
z - Cognitive abilities	0.00786	-0.0342	0.0586	-0.0876	0.0271	0.0471	0.00337	0.04
	(0.0300)	(0.0387)	(0.0464)	(0.0594)	(0.0336)	(0.0473)	(0.0476)	(0.067)
n	1,393	757	636	1,393	1,716	902	814	1,716
z - Behavioural problems	-0.106***	-0.0988**	-0.117**	0.0278	-0.0677*	-0.118**	-0.00289	-0.12
	(0.0322)	(0.0449)	(0.0463)	(0.0634)	(0.0382)	(0.0506)	(0.0582)	(0.077)
n	1,789	995	794	1,789	1,826	982	844	1,826

Notes: FE=OLS fixed effects; robust standard errors in parentheses, controls: birth order for all models and grade track for z-academic performance; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10

Table A.8. Twin FE-models by birth cohort, outcomes, parental education and prematurity (<37 weeks of gestation)

	5-year-old cohort						**11-year-old cohort**					
	Academic Performance											
	All Sample		**Low-Edu.**		**High-Edu.**		**All Sample**		**Low-Edu.**		**High-Edu.**	
	Fullborn	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**
Log(birth weight)	0.150	1.245***	-0.0541	1.400***	0.708*	0.973**	0.661*	0.786***	0.950**	1.035**	0.330	0.629*
	(0.239)	(0.272)	(0.270)	(0.322)	(0.418)	(0.483)	(0.380)	(0.292)	(0.466)	(0.493)	(0.587)	(0.321)
n	792	959	421	560	371	399	648	727	317	381	331	346
	Cognitive Ability											
	All Sample		**Low-Edu.**		**High-Edu.**		**All Sample**		**Low-Edu.**		**High-Edu.**	
	Fullborn	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**
Log(birth weight)	-0.00728	0.203	-0.122	-0.155	0.291	0.699	0.543	-0.0771	0.714	0.202	0.348	-0.494
	(0.202)	(0.273)	(0.239)	(0.348)	(0.493)	(0.439)	(0.375)	(0.272)	(0.501)	(0.361)	(0.567)	(0.402)
n	631	729	330	415	301	314	801	868	400	474	401	394
	Behavioural Problems											
	All Sample		**Low-Edu.**		**High-Edu.**		**All Sample**		**Low-Edu.**		**High-Edu.**	
	Fullborn	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**	**Fullborn**	**Preterm**
Log(birth weight)	-0.735***	-0.703**	-0.587***	-0.664*	-1.170**	-0.764*	-0.0654	-0.650**	-0.208	-0.830**	0.0517	-0.343
	(0.210)	(0.295)	(0.224)	(0.389)	(0.530)	(0.456)	(0.412)	(0.301)	(0.575)	(0.402)	(0.592)	(0.462)
n	792	949	421	556	371	393	840	938	431	523	409	415

Notes: robust standard errors in parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10. Controls: birth order for all models and grade track for z-academic performance.

Table A.9. Twin FE-models by birth cohort and parental education on diagnosis of mental/physical illness

	5-year-old				11-year-old			
	Health Problems							
	All Sample	Low-Edu.	High-Edu.	Low - High	All Sample	Low-Edu.	High-Edu.	Low - High
Log(birth weight)	-0.136	-0.144	-0.112	-0.06	-0.175*	-0.268**	-0.0493	-0.252
	(0.091)	(0.112)	(0.153)	(0.188)	(0.100)	(0.130)	(0.156)	(0.205)
n	1,816	1,008	796	1,816	1,854	990	858	1,854

Notes: robust standard errors in parentheses; two-tailed t-test: ***p-value<0.01; **p-value<0.05; *p-value<0.10. Controls: birth order. Diagnosis = at least one of the following health problems diagnosed: respiratory illness, allergies, neurodermatitis, defective vision, eating disorder, motor dysfunction, mental disability, physical disability, anxiety disorder, social behaviour disorder, attention deficit disorder, dyslexia (reading/writing difficulties), dyscalculia (difficulties with maths), stuttering, other physical or mental illness.

Table A.10. Naïve OLS models by birth cohort, outcomes and parental education

	5-year-old cohort								
	z - Academic Performance			**z - Cognitive Ability**			**z - Behavioural Problems**		
	All Sample	**Low-Edu.**	**High-Edu.**	**All Sample**	**Low-Edu.**	**High-Edu.**	**All Sample**	**Low-Edu.**	**High-Edu.**
Log(birth weight)	0.390**	0.417**	0.366	0.209	0.313	0.145	-0.308**	-0.272	-0.369*
	(0.155)	(0.197)	(0.255)	(0.165)	(0.198)	(0.281)	(0.144)	(0.188)	(0.219)
Mother's age	0.00987	0.0137*	0.00662	0.00794	0.0161**	-0.00303	-0.0198***	-0.0174**	-0.0228***
	(0.00605)	(0.00724)	(0.0108)	(0.00621)	(0.00721)	(0.0122)	(0.00578)	(0.00773)	(0.00816)
Weeks of gestation	0.00466	0.0125	-0.00934	0.0266	0.0213	0.0307	-0.00740	-0.0116	-0.00371
	(0.0167)	(0.0225)	(0.0266)	(0.0174)	(0.0212)	(0.0295)	(0.0157)	(0.0213)	(0.0230)
2nd born twin (1st born)	0.0179	0.0102	0.0303	-0.0597*	-0.0622	-0.0491	-0.0217	-0.0467	0.0105
	(0.0325)	(0.0435)	(0.0493)	(0.0336)	(0.0455)	(0.0504)	(0.0348)	(0.0485)	(0.0497)
Age in months	0.00795	0.00390	0.0112	0.0226**	0.0210*	0.0237	-0.00152	0.00266	-0.00833
	(0.00732)	(0.00989)	(0.0110)	(0.00879)	(0.0111)	(0.0144)	(0.00671)	(0.00970)	(0.00875)
Dizygotic twin-pair (MZ)	0.00415	-0.0142	0.0372	-0.0860	-0.0745	-0.101	0.123**	0.108	0.155**
	(0.0580)	(0.0790)	(0.0878)	(0.0629)	(0.0834)	(0.0966)	(0.0554)	(0.0786)	(0.0764)
Female twin-pair (Male)	0.0502	0.0789	0.0344	0.116*	0.208**	0.0143	-0.125**	-0.125	-0.118
	(0.0567)	(0.0772)	(0.0858)	(0.0623)	(0.0810)	(0.0975)	(0.0550)	(0.0793)	(0.0737)
High parental education (Low)	0.348***			0.165**			-0.187***		
	(0.0588)			(0.0643)			(0.0564)		
Employed parents (Non-employed)	0.00886	-0.0927	0.363*	0.0878	-0.0205	0.361**	0.0935	0.0318	0.320**
	(0.113)	(0.133)	(0.204)	(0.109)	(0.128)	(0.183)	(0.0944)	(0.112)	(0.161)
Migrant parents (Natives)	-0.305***	-0.335***	-0.230	-0.0380	-0.148	0.154	0.106	0.0943	0.141
	(0.0858)	(0.108)	(0.140)	(0.0909)	(0.118)	(0.139)	(0.0809)	(0.102)	(0.133)
Single mothers (Intact family)	-0.196**	-0.209*	-0.178	-0.251***	-0.233**	-0.307***	0.369***	0.374***	0.355**
	(0.0912)	(0.114)	(0.148)	(0.0690)	(0.0930)	(0.0869)	(0.0947)	(0.115)	(0.165)
Constant	-4.273***	-4.540***	-3.718**	-4.441***	-5.288***	-3.689*	3.450***	3.072**	3.835***
	(0.999)	(1.260)	(1.635)	(1.144)	(1.421)	(1.896)	(0.882)	(1.205)	(1.299)
n	1,787	1,000	787	1,504	827	677	1,780	998	782
R-squared	0.080	0.050	0.031	0.051	0.060	0.035	0.065	0.044	0.055

Notes: Clustered standard errors at the family level in parentheses; reference categories in parentheses; ***p<0.01; **p<0.05; *p<0.10

Table A.10. Continued

	11-year-old cohort								
	z - Academic Performance			**z - Cognitive Ability**			**z - Behavioural Problems**		
	All Sample	**Low-Edu.**	**High-Edu.**	**All Sample**	**Low-Edu.**	**High-Edu.**	**All Sample**	**Low-Edu.**	**High-Edu.**
Log(birth weight)	0.206	0.365*	0.0893	0.447***	0.498***	0.423**	-0.250*	-0.481***	0.139
	(0.169)	(0.219)	(0.248)	(0.135)	(0.173)	(0.215)	(0.148)	(0.185)	(0.236)
Mother's age	-0.00143	-0.00642	0.000765	0.0105*	0.00805	0.0121	-0.00710	-0.0144*	0.00523
	(0.00692)	(0.00944)	(0.00991)	(0.00577)	(0.00766)	(0.00839)	(0.00622)	(0.00855)	(0.00884)
Weeks of gestation	-0.00972	-0.0213	-0.00262	-0.00298	-0.00248	-0.00493	-0.000715	0.0102	-0.0202
	(0.0168)	(0.0223)	(0.0248)	(0.0138)	(0.0193)	(0.0198)	(0.0149)	(0.0192)	(0.0232)
2nd born twin (1st born)	-0.0622*	-0.0450	-0.0821*	-0.0364	-0.00470	-0.0671	-0.0178	-0.00227	-0.0288
	(0.0339)	(0.0507)	(0.0454)	(0.0349)	(0.0495)	(0.0493)	(0.0384)	(0.0552)	(0.0531)
Age in months	-0.0410***	-0.0515***	-0.0347***	0.0176**	0.0236**	0.0123	-0.00809	-0.00199	-0.0168*
	(0.00923)	(0.0143)	(0.0118)	(0.00732)	(0.0103)	(0.0101)	(0.00736)	(0.0111)	(0.00922)
Dizygotic twin-pair (MZ)	0.0343	-0.00878	0.0847	-0.0347	-0.140*	0.0886	0.0442	0.123	-0.0918
	(0.0641)	(0.0945)	(0.0852)	(0.0564)	(0.0772)	(0.0809)	(0.0563)	(0.0804)	(0.0786)
Female twin-pair (Male)	0.142**	0.173*	0.0964	0.0919*	0.0290	0.160**	-0.00463	0.0414	-0.0685
	(0.0613)	(0.0919)	(0.0789)	(0.0530)	(0.0755)	(0.0731)	(0.0538)	(0.0794)	(0.0718)
High parental education (Low)	0.465***			0.383***			-0.239***		
	(0.0690)			(0.0572)			(0.0551)		
Employed parents (Non-employed)	0.143	0.306**	-0.269	0.256***	0.489***	-0.471**	0.0265	0.0116	0.0698
	(0.130)	(0.151)	(0.220)	(0.0979)	(0.101)	(0.185)	(0.0995)	(0.119)	(0.153)
Migrant parents (Natives)	-0.334***	-0.323**	-0.391***	-0.166**	-0.132	-0.315**	-0.182**	-0.253***	0.0376
	(0.0944)	(0.125)	(0.138)	(0.0837)	(0.104)	(0.134)	(0.0736)	(0.0929)	(0.122)
Single mothers (Intact family)	-0.184**	-0.126	-0.219	-0.144*	-0.0831	-0.224*	0.197***	0.178*	0.194*
	(0.0851)	(0.107)	(0.140)	(0.0737)	(0.0911)	(0.121)	(0.0738)	(0.0940)	(0.108)
Constant	3.998**	4.542*	4.694**	-6.610***	-7.758***	-4.845**	3.415**	4.087**	1.768
	(1.691)	(2.469)	(2.320)	(1.295)	(1.738)	(1.969)	(1.343)	(1.886)	(1.891)
n	1,488	769	719	1,769	927	842	1,821	964	857
R-squared	0.126	0.072	0.087	0.101	0.071	0.047	0.035	0.035	0.014

Notes: Clustered standard errors at the family level in parentheses; reference categories in parentheses; controls: grade track for z-academic performance. ***p<0.01; **p<0.05; *p<0.1

Table A.11. Main health risk factors during pregnancy (top-10 most prevalent conditions within risk number 1) registered in the health book record by parental SES (column %) for the 11-year-old cohort

Risk Number 1 During Pregnancy	All Sample %	Low-Edu. %	High-Edu. %	Low – High %
Multiple pregnancy	26.04	26.71	25.33	1.38
Over 35 years of age	14.77	11.18	18.56	-7.38
Family history (e.g., diabetes, hypertension, congenital anomalies, genetic disorders, mental illness)	14.35	16.15	12.45	3.7
Allergies, including to medications	11.05	11.59	10.48	1.11
History of fertility treatment	8.29	7.45	9.17	-1.72
Prior severe illnesses, (e.g., heart, lung, liver, kidneys, central nervous system, mental)	4.57	4.55	4.59	-0.04
History of Caesarean section	2.87	2.07	3.71	-1.64
Obesity	2.02	2.90	1.09	1.81
Preterm labour	2.02	1.66	2.40	-0.74
History of 2 or more miscarriages/abortions	1.49	2.48	0.44	2.04
n	941	483	458	

Notes: the health record book was less available among low-educated families (23% missing) than for high-educated families (18% missing). Among those families with an available health book record, the share of unreadable entries stands at 33% (risk number 1) and does not vary by SES.

Table A.12. Naïve OLS models on medical check-ups before giving birth and naïve LPM on LBW (columns 4-5) for the 11-year-old cohort

	Pregnancy Week Initial Examination	**Number of Preventive Examinations**	**Visited Hospital Before Giving Birth**	**LBW**	**LBW**
Mother's age	0.00886	0.0290	-0.00361	-0.000607	-0.000889
	(0.0400)	(0.0299)	(0.00387)	(0.00378)	(0.00377)
Preterm			-0.0483	0.442***	0.439***
			(0.0335)	(0.0346)	(0.0348)
Female twin-pair (Male)	-0.237	0.188	-0.0222	0.0508	0.0490
	(0.337)	(0.279)	(0.0333)	(0.0339)	(0.0339)
Dizygotic-twin pair (MZ)	-0.362	0.403	-0.0312	-0.0396	-0.0420
	(0.352)	(0.296)	(0.0346)	(0.0352)	(0.0351)
High parental education (Low)	-0.388	0.360	0.0756**	-0.0639*	-0.0580
	(0.344)	(0.289)	(0.0349)	(0.0353)	(0.0354)
Visited hospital before giving birth (No)					-0.0781**
					(0.0371)
Constant	9.701***	9.146***	0.979***	0.411**	0.488***
	(1.622)	(1.372)	(0.173)	(0.177)	(0.182)
n	1,391	1,357	1,276	1,276	1,276
R-squared	0.004	0.008	0.012	0.217	0.222

Notes: Clustered standard errors at the family level in parentheses; reference categories in parentheses; LBW: 1=at least one co-twin LBW; 0=both co-twins NBW; ***p<0.01; **p<0.05; *p<0.1

Table A.13. Descriptive statistics and two-tailed t-test by parental education on pre-school/school activities/support for the 11-year-old cohort

Group	n	Mean	Std. Err.	SD
		Pre-school Activities[1]		
Low-edu.	993	3.23	0.04	1.14
High-edu.	883	3.42	0.04	1.06
All sample	1,876	3.32	0.03	1.11
Low-edu. - High-edu.		-0.19***	0.05	
		Pre-school Support[2]		
Low-edu.	987	0.38	0.02	0.48
High-edu.	856	0.31	0.02	0.46
All sample	1,843	0.34	0.01	0.48
Low-edu. - High-edu.		0.07***	0.02	
		School Support[3]		
Low-edu.	987	0.39	0.02	0.49
High-edu.	856	0.31	0.02	0.46
All sample	1,843	0.35	0.01	0.48
Low-edu. - High-edu.		0.07***	0.02	

Notes: ***p-value<0.01; **p-value<0.05; *p-value<0.10; 1=Child taking part during the entire time in childcare up to the age of 6 (normal kindergarten and Kita activities as well as special activities beyond the standard program) in the following activities: early musical education, drawing/painting, help with learning German, foreign language classes, mathematical and scientific stimuli, visits to libraries and nature trips. From these items, we estimated an average index in the original scale ranging from 0 (the activity was not offered) to 6 (several times a week). 2=Twin-specific use of the support provided by kindergarten or another institution for special needs education in one or more of the following: learning, speaking, physical and motor development, emotional and social development, mental development, vision, hearing, autism and/or others. 3=Twin-specific participation outside regular school hours in one or more of the following activities: help with homework; remedial groups; and/or subject-specific additional classes.

Table A.14. FE-OLS models on pre-school and school support for the 11-year-old cohort

	Pre-school Support[1]				School Support[2]			
	All Sample	**Low-Edu.**	**High-Edu.**	**Low - High**	**All Sample**	**Low-Edu.**	**High-Edu.**	**Low - High**
Log(birth weight)	-0.255***	-0.279**	-0.223	-0.072	0.151**	0.0887	0.237**	-0.139
	(0.0902)	(0.116)	(0.142)	(0.181)	(0.0696)	(0.0972)	(0.0991)	(0.140)
Observations	1,843	987	856	1,843	1,843	987	856	1,843

Notes: Robust standard errors in parentheses; two-tailed t-test: Controls: birth order; ***p-value<0.01; **p-value<0.05; *p-value<0.10;
1=Twin-specific use of the support provided by kindergarten or another institution for special needs education in one or more of the following: learning, speaking, physical and motor development, emotional and social development, mental development, vision, hearing, autism and/or others.
2=Twin-specific participation outside regular school hours in one or more of the following activities: help with homework; remedial groups; and/or subject-specific additional classes.

Table A.15. Mean and two-tailed t-test by variables and family-average parental investments' quintiles for the 11-year old cohort

Variables	**Family-mean Cultural Activities**			**Family-mean Parental Warmth**		
	Low (q1)	**High (q2-q5)**	**Low - High**	**Low (q1)**	**High (q2-q5)**	**Low – High**
Mother's age	42.15	43.19	-1.04***	42.54	43.24	-0.71***
High parental education	0.33	0.50	-0.16***	0.40	0.51	-0.11***
Highest parental ISEI	46.48	54.61	-8.13***	48.55	55.75	-7.20***
z - Within-family differences in cultural activities	0.46	0.90	-0.44***			
z - Within-family diff. in parental warmth				1.03	0.68	0.35***
Dizygotic twin-pair	0.55	0.61	-0.06**	0.57	0.60	-0.03
Female twin-pair	0.44	0.54	-0.10***	0.52	0.52	0.00
Birth weight (grams)	2356.74	2437.46	-80.72***	2398.81	2434.16	-35.35
Within-family birth weight differences (grams)	271.25	268.01	3.24	256.46	272.83	-16.37
LBW	0.55	0.50	0.05*	0.55	0.49	0.06*
Gestational weeks	35.52	35.93	-0.41***	35.66	35.91	-0.25*
Preterm	0.56	0.52	0.04	0.53	0.53	0.01
z – Foetal growth	-0.09	0.03	-0.12**	-0.01	0.02	-0.03
Diagnosis	0.43	0.45	-0.03	0.45	0.45	0.00
Pregnancy week initial examination	9.70	8.78	0.92***	9.41	8.78	0.64**
Visited hospital before giving birth	0.76	0.75	0.01	0.74	0.76	-0.02
Number of preventive examinations	11.05	11.66	-0.61**	11.03	11.66	-0.63**
n	250-406	1,080-1,618		251-398	1,063-1,518	

Notes: ***p-value<0.01; **p-value<0.05; *p-value<0.10

Table A.16. Distribution of key variables in twin and singleton samples

Population Sample					
National Educational Panel Study SC2 Refreshment Sample					
Year of birth 2005/2006 – Wave 3 2012/2013					
	Mean	**SD**	**Min.**	**Max.**	**Obs.**
Birth weight (grams)	3,347.42	589.97	302	7,800	5,022
Low birth weight (<2,500gr.)	6.3%		0	100	5,022
Preterm births (<37 weeks)	9%+		0	100	n/a
Mother's age at child birth	29.3	5.33	13	54	5,022
High parental education*	36%		0	100	5,022
Twin Sample					
Twin Life Study Sample					
Year of birth 2003/2004 – Wave 1 2014/2016					
	Mean	**SD**	**Min.**	**Max.**	**Obs.**
Birth weight (grams)	2,420.12	538.97	630	4,300	1,937
Low birth weight (<2,500gr.)	51.1%		0	100	1,937
Preterm births (<37 weeks)	53%		0	100	1,937
Mother's age at child birth	32.1	4.91	17	47	1,937
High parental education*	46%		0	100	1,937

+*Source:* Euro-Peristat Project (2008). *European Perinatal Health Report. Core indicators of the health and care of pregnant women and babies in Europe in 2004.* www.europeristat.com; *Highest parental education = ISCED 5A+6

References

Aarnoudse-Moens, Cornelieke S.H., et al. 2009. “Meta-Analysis of Neurobehavioural Outcomes in Very Preterm and/or Very Low Birth Weight Children.” *Pediatrics, 124*(2):717LP–728.

Abernethy, Laurence J., Mike Palaniappan, and Richard W.I. Cooke. 2002. “Quantitative magnetic resonance imaging of the brain in survivors of very low birth weight.” *Archives of Disease in Childhood*, 87:279-283.

Abufhele, Alejandra, Jere Behrman, and David Bravo. 2017. “Parental preferences and allocations of investments in children’s learning and health within families.” *Social Science and Medicine*, *194*:76–86.

Aizer, A., Stroud, L., Buka, S. 2016. Maternal Stress and Child Outcomes Evidence from Siblings. *Journal of Human Resources, 51*(3), 523-555. doi:10.3368/jhr.51.3.0914-6664R

Aizer, Anna, and Flávio Cunha. 2012. "The Production of Human Capital: Endowments, Investments and Fertility." *NBER Working Papers 18429, National Bureau of Economic Research, Inc.*

Aizer, Anna, and Janet Currie. 2014. The intergenerational transmission of inequality: Maternal disadvantage and health at birth. *Science*, *344*(6186), 856 LP – 861.

Almond, D. 2006. Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. Journal of Political Economy, 114(4), 672-712. doi:10.1086/507154

Almond, D., Edlund, L., and Palme, M. 2009. Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. Quarterly Journal of Economics, 124(4), 1729-1772.

Almond, Douglas, and Bhashkar Mazumder. 2013. “Fetal Origins and Parental Responses.” *Annual Review of Economics*, 5:37-56.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature*, 56 (4):1360-1446.

Asbury, Kathryn, Judith F. Dunn, and Robert Plomin. 2006. “Birth weight-discordance and differences in early parenting relate to monozygotic twin differences in behaviour problems and academic achievement at age 7.” *Developmental Science*, *9*(2):F22–F31.

Attanasio, Orazio, et al. 2020. "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia." *American Economic Review,* 110 (1): 48-85.

Baranowska-Rataj, A., Barclay, K., Costa-Font, J., Myrskylä, M. and Özcan, B. 2019. "Preterm Births and Educational Disadvantage: Heterogeneous Effects across families and Schools." *Social Policy Working Paper* 09-19, London: LSE Department of Social Policy.

Barker, David J.P. 1990. "The fetal and infant origins of adult disease." *BMJ: British Medical Journal*, 301(6761):1111.

Becker, Gary and Nigel Tomes. 1976. "Child endowments and the quantity and quality of children." *Journal of Political Economy*, 84(4), part 2.

Behrman, Jere R., Robert A. Pollak, and Paul Taubman. 1982. "Parental Preferences and Provisions for Progeny." *Journal of Political Economy*, 90(1):52.

Bernardi, Fabrizio. 2014. "Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth." *Sociology of Education*, 87(2):74–88.

Bharadwaj, Prashant, Juan P. Eberhard, and Christopher A. Neilson. 2017. "Health at Birth, Parental Investments, and Academic Outcomes." *Journal of Labor Economics, 36*(2):349–394.

Black, S. E., Devereux, P. J., and Salvanes, K. G. 2016. Does grief transfer across generations? bereavements during pregnancy and child outcomes. American Economic Journal: Applied Economics, 8(1), 193-223.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the cradle to the labor market? The effect of birth weight on adult outcomes." *Quarterly Journal of Economics, 122*(1):409-439.

Blencowe, Hannah, et al. 2019. "National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis." *The Lancet Global Health*, 7(7):PE849-E860

Boardman, Jason D., et al. 2002. "Low Birth Weight, Social Factors, and Developmental Outcomes among Children in the United States." *Demography*, 39(2):353–368.

Brix, J., Pupeter, M., Rysina, A., Steinacker, G., Schneekloth, U., Baier, T., Gottschling, J., [...] & Spinath, F. M. 2017. "A longitudinal twin family study of the life course and individual development (TwinLife): Data collection and instruments of wave 1 face-to-face interviews." *TwinLife Technical Report Series, 05. Project TwinLife: Genetic and social causes of life chances* (Universität Bielefeld / Universität des Saarlandes).

Brown, R. 2018. The Mexican drug war and early-life health: The impact of violent crime on birth outcomes. *Demography*, *55*(1), 319-340.

Camacho, A. 2008. Stress and birth weight: evidence from terrorist attacks. *The American Economic Review*, 98(2), 511-515.

Cleary-Goldman, J., and Alton, M. E. D. 2008. Growth Abnormalities and Multiple Gestations, 206–212. *Seminars in Perinatology*, 32(3): 206-212.

Conley, D., Rauscher, E. 2012. "Genetic Interactions with Prenatal Social Environment: Effects on Academic and Behavioural Outcomes." *Journal of Health and Social Behaviour*, 54(1):109–127

Conley, D., Sotoudeh, R., Laidley, T. 2019. Birth Weight and Development: Bias or Heterogeneity by Polygenic Risk Factors? *Population Research and Policy Review*, *38*(6), 811–839. https://doi.org/10.1007/s11113-019-09559-6

Conley, Dalton, Kathryn M. Pfeiffer, and Melissa Velez. 2007. "Explaining sibling differences in achievement and behavioural outcomes: The importance of within- and between-family factors." *Social Science Research*, 36(3):1087–1104.

Conley, Dalton, Neil G. Bennett. 2000. "Is biology destiny? Birth weight and life chances." *American Sociological Review*, 65(3):458-467.

Conley, Dalton. 2008. "Bringing Sibling Differences In: Enlarging Our Understanding of the Transmission of Advantage in Families." In Conley D. Lareau A. (Eds.), *Social Class: How Does It Work?* (pp. 179-200). Russell Sage Foundation.

Conti, Gabriella, et al. 2018. "Beyond Birthweight: The Origins of Human Capital." *Institute for Fiscal Studies Working Paper*:18-30.

Currie, J., and Rossin-Slater, M. 2013. Weathering the storm: Hurricanes and birth outcomes. Journal of health economics, 32(3), 487-503.

Currie, J., and Schwandt, H. 2015. The 9/11 Dust Cloud and Pregnancy Outcomes: A Reconsideration. *Journal of Human Resources.*

Currie, Janet and Douglas Almond. 2011. "Chapter 15 - Human capital development before age five." In: D. Card and O. Ashenfelter (Eds.): *Handbook of Labor Economics* (Vol. 4, pp. 1315–1486). Elsevier.

Currie, Janet and Rosemary Hyson 1999. "Is the Impact of Health Shocks Cushioned by Socio-Economic Status? The Case of Birth Weight." *American Economic Review*, 89(2):245-250.

Currie, Janet. 2011. "Inequality at Birth: Some Causes and Consequences." *American Economic Review, 101*(3): 1-22.

Datar, Ashlesha, M. Rebecca Kilburn, and David S. Loughran. 2010. "Endowments and parental investments in infancy and early childhood." *Demography*, *47*(1), 145–162.

Diewald, Martin et al. 2018. *TwinLife. GESIS Data Archive, Cologne. ZA6701 Data file Version 2.1.0*, doi:10.4232/1.13072.

Duncan, B., Mansour, H., and Rees, D. I. 2017. It's Just a Game: The Super Bowl and Low Birth Weight. *Journal of Human Resources*, *52*(4), 946-978.

Farah, Martha J., et al. 2008. "Environmental stimulation, parental nurturance and cognitive development in humans." *Developmental Science*, 11:793-801

Figlio, David, et al. 2014. "The Effects of Poor Neonatal Health on Children's Cognitive Development." *American Economic Review, 104*(12):3921–3955.

Francesconi, Marco, and James J. Heckman. 2016. "Child Development and Parental Investment: Introduction." *The Economic Journal*, *126*(596):F1–F27.

Gielen, M., Lindsey, P. J., Derom, C., Smeets, H. J. M., Souren, N. Y., Paulussen, A. D. C., … Nijhuis, J. G. 2008. Modeling genetic and environmental factors to increase heritability and ease the identification of candidate genes for birth weight: A twin study. *Behaviour Genetics*, *38*(1), 44–54. https://doi.org/10.1007/s10519-007-9170-3

Goosby, Bridget J., and Jacob E. Cheadle. 2009. "Birth Weight, Math and Reading Achievement Growth: A Multilevel Between-Sibling, Between-Families Approach." *Social Forces*, *87*(3):1291–1320.

Grätz, Michael, and Florencia Torche. 2016. "Compensation or Reinforcement? The Stratification of Parental Responses to Children's Early Ability." *Demography*, 53:1883-1904.

Grow, Can Well-Educated Parents Buffer This Adverse Factor? A Research Note." *Demography* 44(2):335-43.

Härkönen, J., Lindberg, M., Karlsson, L., Karlsson, H., & Scheinin, N. M. 2018. Education is the strongest socioeconomic predictor of smoking in pregnancy. *Addiction.*

Heckman, James. J. 2007. "The economics, technology, and neuroscience of human capability formation." *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33):13250-13255.

Hsin, Amy 2012. "Is biology destiny? Birth weight and differential parental treatment." *Demography*, *49*(4):1385-1405.

Isen, A., Rossin-Slater, M., and Walker, W. R. 2017. Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970. Journal of Political Economy, 125(3), 848-902.

Jæger, Mads M. 2011. "Does Cultural Capital Really Affect Academic Achievement? New Evidence from Combined Sibling and Panel Data." *Sociology of Education* 84(4):281–98.

Kelly, Yvonne J., et al. 2001. "Birthweight and behavioural problems in children: a modifiable effect?" *International Journal of Epidemology* 30(1):88-94

Kim, B., Carruthers, C. K., and Harris, M. C. 2017. Maternal stress and birth outcomes: Evidence from the 1994 Northridge earthquake. Journal of Economic Behavior & Organization, 140, 354-373.

Kim, Soo Hyun, Yoo-Min Kim, Ji-Hee Sung, Suk-Joo Choi, Soo-Young Oh & Cheong-Rae Roh (2020). The effects of birth order on neonatal outcomes in early-preterm, late-preterm and term twin infants, *The Journal of Maternal-Fetal & Neonatal Medicine*, 33:12, 1980-1987.

Knopik, Valerie S., et al. 2017. *Behavioural genetics* (7th ed.). NY: Worth.

Knudsen, Eric I., et al. 2006. "Economic, neurobiological, and behavioural perspectives on building America´s future workforce." *Proceedings of the National Academy of Sciences*, 27:10155-10162.

Kohler, Hans-Peter, Behrman, Jere R., and Jason Schnittker. 2011. "Social science methods for twins data: Integrating causality, endowments, and heritability." *Biodemography and Social Biology*, 57(1):88–141.

Lang, Volker, and Anita Kottwitz. 2017. *The sampling design and socio-demographic structure of the first wave of the TwinLife panel study: a comparison with the Microcensus* (TwinLife Technical Report Series, 03) updated version, August 2017. Bielefeld: Project TwinLife "Genetic and social causes of life chances".

Lareau, Annette. 2003. *Unequal Childhoods: Class, Race, and Family Life.* Berkeley and Los Angeles: University of California Press.

Laucht, Manfred, Günter Esser, and Martin Schmidt. 2001. "Developmental outcome of infants born with biological and psychosocial risks." *Journal of Child Psychiatry*, 38:843–853.

Layte, Richard. 2017. "Why Do Working-Class Kids Do Worse in School? An Empirical Test of Two Theories of Educational Disadvantage." *European Sociological Review*, 33(4):489-503.

Lesaar, S., Prussog-Wagner, A., and Hess, D. 2020. "TwinLife survey methodology and fieldwork outcomes. Face-to-face survey of wave 2 (F2F 2a/b)." *TwinLife Technical Report*

Series; 10, v1.0.0. Bielefeld: Project TwinLife "Genetic and social causes of life chances" (Universität Bielefeld / Universität des Saarlandes).

Lin, Ming-Jen, Jin-Tan Liu and Shin-Yi Chou. 2007. "As Low Birth Weight Babies Grow, Can Well-Educated Parents Buffer This Adverse Factor? A Research Note." *Demography* 44(2):335-43.

Lugo-Gil, Julieta, and Catherine S. Tamis-LeMonda. 2008. "Family resources and parenting quality: Links to children's cognitive development across the first 3 years." *Child Development*, 79(4):1065-1085.

Lynch, Jamie L., and Benjamin G. Gibbs. 2016. "Birth Weight and Early Cognitive Skills: Can Parenting Offset the Link?" *Maternal and Child Health Journal*, 21(1):156-167.

McCormick, Marie C., et al. 2006. "Early Intervention in Low Birth Weight Premature Infants: Results at 18 Years of Age for the Infant Health and Development Program." *Pediatrics*, *117*(3):771LP –780.

McEwen, Craig A., and Bruce S. McEwen. 2017. "Social Structure, Adversity, Toxic Stress, and Intergenerational Poverty: An Early Childhood Model." *Annual Review of Sociology*, 43:1–28.

McGue, Mat, Merete Osler, and Kaare Christensen. 2010. "Causal Inference and Observational Research: The Utility of Twins." *Perspectives on Psychological Science*, *5*(5):546–556

McLean, M., Bisits, A., Davies, J., Woods, R., Lowry, P., & Smith, R. (1995). A placental clock controlling the length of human pregnancy. *Nature medicine*, *1*(5), 460-463.

Møllegaard, S. 2020. The effect of birth weight on behavioural problems in early adolescence: New evidence from monozygotic twins. *Economics and Human Biology*, *36*. https://doi.org/10.1016/j.ehb.2019.100828

Mönkediek, B., Schulz, W., Eichhorn, H., & Diewald, M. 2020. Is there something special about twin families? A comparison of parenting styles in twin and non-twin families. *Social Science Research*, *90*, 102441.

Newcombe, R., Milne, B., Caspi, A., Poulton, R., & Moffitt, T. 2007. Birthweight Predicts IQ: Fact or Artefact? *Twin Research and Human Genetics*, *10*(4), 581-586.

Oreopoulos, Philip, et al. 2008. "Short, medium, and long-term consequences of poor infant health: An anlysis using siblings and twins." *Journal of Human Resources*, 43(1):88-138.

Pampel, F. C., Krueger, P. M., & Denney, J. T. (2010). Socioeconomic disparities in health behaviors. *Annual Review of Sociology*, *36*, 349.

Persson, P., and Rossin-Slater, M. 2018. Family ruptures, stress, and the mental health of the next generation. *American Economic Review, 108*(4-5), 1214-1252.

Raisanen, S., Gissler, M., Sankilampi, U., Saari, J., Kramer, M. R., & Heinonen, S. (2013). Contribution of socioeconomic status to the risk of small for gestational age infants - a population-based study of 1,390,165 singleton live births in Finland. *International Journal for Equity in Health, 12*, 8. doi:10.1186/1475-9276-12-28

Ranson, Kenna and Liana Urichuk (2008): "The Effect of Parent-Child Attachment Relationships on Child Biopsychosocial Outcomes: A Review," *Early Child Development and Care*, 178(2):129-152.

Restrepo, Brandon J. 2016. "Parental investment responses to a low birth weight outcome: who compensates and who reinforces?" *Journal of Population Economics*, 29(4):969–989.

Skopek, Jan, and Giampiero Passaretta. 2018. "The Social Stratification of Skills from Infancy to Adolescence: Evidence from an Accelerated Longitudinal Design." *SocArXiv. October 1.*

Slama, R., Darrow, L., Parker, J., Woodruff, T., Strickland, M., Nieuwenhuijsen, M., . . . Hurley, F. (2008). Atmospheric pollution and human reproduction: report of the Munich International Workshop. *Environ Health Perspect, 116*(6), 791-798.

Smithers, Lisa G., et al. 2018. "A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes." *Nature Human Behaviour, 2*(11):867–880.

Sripada, Rebecca K., et al. 2018. "Trajectories of brain development in school-age children born preterm with very low birth weight." *Nature Scientific Reports, 8*(1):15553.

Susser, M., and Stein, Z. (1994). Timing in prenatal nutrition - a reprise of the dutch famine study. Nutrition Reviews, 52(3), 84-94.

Torche, F., and Villarreal, A. (2014). Prenatal exposure to violence and birth weight in Mexico selectivity, exposure, and behavioral responses. American Sociological Review, 79(5), 966-992.

Torche, Florencia, and Dalton Conley. 2016. "A Pound of Flesh: The Use of Birth weight as a Measure of Human Capital Endowment in Economics Research." In J. Komlos and I. Kelly (Eds.), *The Oxford Handbook of Economics and Human Biology*, New York, NY: Oxford University Press.

Torche, Florencia, and Echevarria, G. 2011. "The effect of birthweight on childhood cognitive development in a middle-income country." *International Journal of Epidemiology, 40*(4):1008-1018.

Torche, Florencia. 2018. "Prenatal Exposure to an Acute Stressor and Children's Cognitive Outcomes." *Demography*, *55*(5):1611–1639.

Tully, Lucy A., et al. 2004. "Does Maternal Warmth Moderate the Effects of Birth Weight on Twins' Attention-Deficit/Hyperactivity Disorder (ADHD) Symptoms and Low IQ?" *Journal of Consulting and Clinical Psychology*, *72*(2):218-226.

van Ewijk, R. 2011. Long-term health effects on the next generation of Ramadan fasting during pregnancy. Journal of health economics, 30(6), 1246-1260. doi:10.1016/j.jhealeco.2011.07.014

Victoria, A., G., Mora, and F., Arias. 2001. Perinatal outcome, placental pathology, and severity of discordance in monochorionic and dichorionic twins. *Obstetrics and Gynecology*, *97*(2), 310–315.

Walhovd, Kristine B., et al. .2012. "Long-term influence of normal variation in neonatal characteristics on human brain development." *Proceedings of the National Academy of Sciences*, 109(49):20089-20094

Walker, Susan P., et al. 2010. "The effect of psychosocial stimulation on cognition and behaviour at 6 years in a cohort of term, low-birthweight Jamaican children." *Developmental Medicine Child Neurology*, *52*(7):e148–e154.

Yi, Junjian, et al. 2015. "Early Health Shocks, Intra-household Resource Allocation and Child Outcomes." *The Economic Journal*, *125*(588):347–371.

Yokoyama, Y., Jelenkovic, A., Sund, R., Sung, J., Hopper, J. L., Ooki, S., … Tarnoki, D. L. (2016). Twin's Birth-Order Differences in Height and Body Mass Index from Birth to Old Age: A Pooled Study of 26 Twin Cohorts Participating in the CODA Twins Project. *Twin Research and Human Genetics, 19(02), 112–124.*

Chapter III

Do Low-IQ but Advantaged Kids Get Ahead? A Twin Study on Early Schooling Inequalities

Carlos J. Gil-Hernández

Abstract

This article bridges the literature on educational inequality between and within families to test whether high–socioeconomic status (SES) families compensate for low cognitive ability in the transition to secondary education in Germany. The German educational system of early-ability tracking (at age 10) represents a stringent setting for the compensatory hypothesis. Overall, previous literature offers inconclusive findings. Previous research between families suffers from the misspecification of parental SES and ability, while most within-family research did not stratify the analysis by SES or the ability distribution. To address these issues, I draw from the *Twin Life* study to implement a twin fixed-effects design that minimises unobserved confounding. I report two main findings. First, highly educated families do not compensate for twins' differences in cognitive ability at the bottom of the ability distribution. Second, holding parents' and children's cognitive ability constant, pupils from highly educated families are 27% more likely to attend the academic track. This result implies a wastage of academic potential for disadvantaged families, challenging the role of cognitive ability as the leading criterion of merit for liberal theories of equal opportunity. These findings point to the importance of other factors that vary between families with different resources such as non-cognitive abilities, risk aversion to downward mobility, and teachers' bias, thus explaining educational success.

1. Introduction

Cognitive ability[45] is one of the strongest predictors of learning and educational outcomes (Deary and Johnson 2010; Deary et al. 2007). However, cognitive ability does not necessarily lead to future educational achievement. Disadvantaged children with similar academic potential or ability as advantaged children have systematically fewer chances for educational success (Bukodi, Erikson and Goldthorpe 2014; Jackson 2013; Papageorge and Thom 2018). Research on inequality *between families* argues that this gap is particularly large among children with low scholastic ability because affluent parents tend to use compensatory strategies to offset the effect of low ability (Bernardi 2014; Esping-Andersen and Cimentada 2018).

Do similar compensation mechanisms also work *within families*? Building on the classic microeconomics literature on the intra-household allocation of resources (Behrman, Pollak and Taubman 1982), Conley (2008) hypothesises that high-socioeconomic status (SES) families are more prone to compensate for siblings' differences in endowments/traits. Due to their extensive pool of cultural and economic resources, siblings with lower ability in high-SES families can reach the same educational outcomes as their more gifted siblings.

Previous studies offer inconclusive and limited evidence on the compensatory hypothesis for four reasons. First, between-family models misspecify the total effect of social background (underestimation) and academic ability (overestimation) due to unobserved heterogeneity (Jæger 2011). Second, between-family models assume, by design, that siblings achieve the same educational outcomes, but this is not necessarily the case (Conley 2004). Third, within-family research has focused on the effect of birth weight (BW) on educational outcomes as a proxy for early ability, rather than using more direct measures of academic ability (Grätz and Torche 2016). Fourth, most previous studies using sibling/twin fixed-effects (FE) estimators do not stratify analyses by parental SES or across the absolute endowment distribution. Thus, whether high-SES families compensate for children's low academic ability is an open empirical question.

To address these limitations of previous research, I use twin fixed-effects to test whether high-SES parents compensate for children's low cognitive ability in the transition to

[45] I use the terms general cognitive abilities, cognitive abilities, intelligence, IQ, and academic ability as equivalents for the sake of simplicity.

secondary education. Twins are born under the same parental circumstances and share at least 50% of their genetic makeup, thus ruling out most sources of confounding in between-family and sibling models. I also look at the heterogeneity of the effect of cognitive ability on track choice across parental SES and the absolute ability distribution. The compensatory hypothesis should be tested at the bottom of the ability distribution.

This article makes the two-fold contribution of bridging the literature on between- and within-family inequality and answering two novel research questions: (1) how does parental SES moderate the effect of within-family differences in cognitive abilities on track choice? (2) how do high-SES families compensate for within-family differences in cognitive abilities at the bottom of the absolute ability distribution?

To answer these questions, I use the first wave of the *Twin Life* study carried out in 2014/2015 (Hahn et al. 2016): a representative survey of the German population with a sample of same-sex 11-year-old twins at grades 5 and 6. The German educational system of early-ability tracking is an interesting scenario for testing the compensatory hypothesis. Because teachers are supposed to recommend tracks on the basis of observed performance, parents may have less discretion to influence track decisions for their low academic-ability children. Hence, the German situation provides a stringent test of the compensatory hypothesis in comparison to educational systems without early tracking (Conley and Glauber 2008).

Results show that highly-educated families do not compensate for children's low academic ability at the bottom of the ability distribution. However, highly-educated families still have substantially larger transition rates to the academic track, even when controlling for parents' and children's cognitive abilities. These findings point to the importance of other unobserved factors that vary between families and could explain educational success (e.g., non-cognitive skills, risk aversion, and teachers' bias). I carried out several robustness checks on reverse causality, confounding, attenuation and sample selection bias, and moderation and found consistent results.

2. Theoretical Background

2.1. Cognitive Ability and Educational Outcomes

Intelligence or general cognitive ability (*g*) is a theoretical construct and a highly valid and reliable measure (Nisbett et al. 2012). After adolescence, it is one of the most stable behavioural traits. It represents a latent factor of several sub-dimensions of analytical abilities, such as verbal, spatial, reasoning, perceptual speed, and working memory that are highly and positively inter-correlated and less genetically influenced than the general construct (Knopik et al. 2017).

Cognitive abilities have been traditionally considered as an innate, productivity-enhancing capacity rewarded in both educational systems and labour markets (Fishkin 2014). Some authors have justified intergenerational inequality and legitimated the stratification system due to its genetic heritability (Jensen, 1969; Hernstein and Murray, 1994). However, cognitive ability is not fixed at birth or genetically determined, but it is malleable and dependent on environmental quality (Farah et al. 2008; Gottschling et al. 2019; Guo and Stearns 2002; Kendler et al. 2015; Ritchie and Tucker-Drob 2018; Tucker-Drob, Briley and Harden 2013).

Intelligence can be divided into two components: analytic or fluid (non-verbal), and crystallised (i.e., store of knowledge, vocabulary, and arithmetic operations). Fluid or non-verbal intelligence tests (e.g., the Raven test) measure a person's capacity to reason and solve novel problems, independent of previous knowledge. Fluid intelligence quotient (IQ) tests are less influenced by sociocultural factors, but they cannot be considered completely context-free. Fluid IQ tests do not directly measure creativity, knowledge, social sensitivity, or domain-specific cognitive competencies (e.g., reading, mathematics, or scientific literacy) that are the target of large-scale international assessment studies such as PISA, TIMSS, and PIAAC (Weinert et al. 2011).

Not surprisingly, IQ is a good predictor of educational performance and competencies (i.e., standardised tests: the American SAT, the British GCSE, or PISA), with correlations ranging from 0.4 to 0.7 (Deary et al. 2007; Erikson and Rudolphi 2010; Neisser et al. 1996; Rindermann 2007). This means up to 50% of the variance would be accounted for by intelligence, leaving ample room for other psychological or non-cognitive characteristics to play a role, such as motivation and perseverance (Almlund et al. 2011; Poropat 2009).

Cognitive ability is less related to school or teacher-assigned grades than to educational competencies, as grades are more influenced by non-cognitive or behavioural factors (Duckworth, Quinn and Tsukayama 2012). This research focuses on fluid IQ as a proxy for academic ability or potential (Erikson and Rudolphi 2010), which is less subject to sociocultural context and individual effort compared to measures of academic performance (Bailey et al. 2017).

2.2. Between-Family Inequality

Between-family models evaluate educational inequalities by drawing a random individual from each family, typically using formalised, rational action theories (RAT) (Breen and Goldthorpe 1997). Rational action theories are built on the formal decomposition of the association between social background and educational attainment into two effects: primary and secondary (Boudon 1974). Primary effects denote the systematic association between parental SES and children's academic performance, which is shaped by genetic,[46] psychological, and cultural factors (Jackson 2013). Secondary effects account for upper-class children's advantage in transition rates to higher educational levels compared to their working-class counterparts when controlling for performance.

Goldthorpe (2007) points to three plausible reasons why working-class children with similar academic performance as more advantaged children would systematically follow less ambitious educational pathways, or be more prone to dropout: (1) relative risk aversion (RRA), i.e., to avoid downward social mobility or demotion, lower educational outcomes suffice (a floor effect); (2) economic resources to afford the direct (e.g., tuition fees), indirect (e.g., living costs), and opportunity (e.g., forgone earnings) costs to keep studying are less available and less stable;[47] and (3) lower actual and perceived chances of success due to poorer average performance, along with underestimated or conservative perceived benefits of education.

2.2.1. The Compensatory Advantage Mechanism

Most research following the rational action framework assumes that differences in transition rates between working-class and advantaged children remains constant across the academic-ability distribution (Jackson 2013). In other words, the lion's share of differences in transition rates by social background would be found in the middle of the academic-ability distribution. The primary rationale is that disadvantaged families have more difficulty evaluating the chances of success in the next educational level when their children are just below or above academic thresholds (Bernardi and Cebolla-Boado 2014).

[46] Recent research has found that the genetic variants (i.e., polygenic scores) associated with educational attainment are almost equally distributed among children in low- and high-income families in the United States (Papageorge and Thom 2018).
[47] These direct and indirect costs of studying are trivial for the German case, where secondary education is free (Stocké 2007).

Alternatively, some authors argue that social inequality in transition rates tends to concentrate among low-performing children from advantaged families, who enjoy larger transition rates to upper secondary school (Bernardi and Triventi 2018). Namely, upper-class families actively compensate for bad or mediocre academic performance to avoid intergenerational downward mobility (e.g., through private tutoring; parental involvement with homework; and expectations). Thus, the central point of the compensatory advantage mechanism is that the life-course trajectories of students from privileged backgrounds are less dependent on prior negative outcomes or disadvantageous traits (Bernardi 2014; Erola and Kilpi-Jakonen 2017).

2.2.2. Previous Findings

Half a century ago, Sewell and Shah (1968) showed that 58% of United States children with a low IQ and highly-educated parents attended college, whereas only 9.3% of children with a low IQ from low-educated families did the same. These differences were relatively constant across the middle (78.9 and 22.9%) and top (91.1 and 40.1%) IQ tertiles. To my knowledge, Bukodi and colleagues (Bukodi et al. 2014; Bukodi, Bourne and Betthäuser 2017) provide the only recent evidence on the interaction between cognitive abilities and parental background in the transition to upper secondary in Britain and Sweden. They did not find a clear moderation effect of parental background in Sweden; in England, they found inequalities concentrated among pupils in the top cognitive quintile.

2.3. Within-Family Inequality

Compared to between-family models, studying inequality dynamics within families can account for a larger array of characteristics shared by siblings (i.e., neighbourhood, school, genes, and parental environment) (Conley et al. 2015; McGue, Osler and Christensen 2010). Namely, by drawing random individuals from different families, we cannot control for factors that siblings share or do not share (Turkheimer and Harden 2014). The environmental factors that siblings share contribute to their overall social background (Sieben and De Graaf 2003) but several factors can vary between siblings in the same family: (1) parental circumstances may affect siblings in diverse ways (e.g., mother's age, birth spacing and order, and shocks such as divorce or employment loss) (Grätz 2018; Härkönen 2014), and (2) extrinsic (e.g., luck, random events) and intrinsic (e.g., unique traits) elements that are specific to each sibling.

Siblings share only about 50% of their genetic makeup, on average, and they have unique environmental experiences (e.g., teachers, friends) that are associated with personality traits (e.g., active self-selection). Accordingly, some authors claim that about 65% of the variation in early academic performance (i.e., reading and mathematics), and around half of the variation in educational attainment, is observed within families in the United States (Conley 2008; Conley, Pfeiffer and Velez 2007) and Germany (Grätz 2018). These unique endowments and personality traits may also evoke different parental treatments or responses (Tucker-Drob et al. 2013): parents may consciously or unconsciously behave in a neutral way, compensate for, or reinforce siblings' initial differences in traits associated with early educational outcomes.

2.3.1. Parental Response to Child Endowments

Most theoretical contributions and findings on within-family inequality come from microeconomic models on intra-household resource allocation (e.g., child-specific investment in human capital) as a function of children's endowments (Becker and Tomes 1976).

The family wealth model (Becker and Tomes 1976) posits that, under the assumption of no capital constraints, parents try to maximise returns on human capital investment by either investing equally/neutrally in both children (over time, *initial* ability differences will unfold, thus generating reinforcement patterns) or investing more in the higher-ability child, thus reinforcing sibling differences in endowments. Alternatively, the separable earnings-transfer model (Behrman and colleagues 1982), based on parental preferences and (within-family) inequality aversion, hypothesises that parents tend to compensate for sibling differences in endowments by investing more in the lower-ability child to maximise their children's human capital and earnings. Overall, patterns of reinforcement for early endowments are commonly found for educational investments in comparison to health investments (Yi et al. 2015).

2.3.2. Within-Family (In)equality by Parental SES: Compensation or Reinforcement?

The literature discussed so far offers limited and mixed findings due to different research designs and measures of early ability or developmental potential (e.g., most research focuses on BW). Furthermore, these analyses do not stratify by parental SES, because within-family (in)equalities may depend on families' pools of resources (Lynch and Brooks 2013).

Conley (2008) builds on the microeconomics literature to theorise about different patterns of within-family inequality by parental SES. In a similar vein to the compensatory advantage mechanism, he suggests that wealthy families, thanks to their reliance on a large pool of cultural and economic resources, tend to *compensate* for within-family differences in endowments. In such cases, children with less academic ability will achieve the same results as their more endowed siblings, generating within-family equality. By contrast, disadvantaged families, due to a scarcity of resources, tend to behave more *efficiently* by "betting" on the sibling with more academic potential, thus reinforcing within-family inequality.

An alternative hypothesis (Becker and Tomes 1986) posits that, in the event of capital constraints, disadvantaged families "may not be able to optimally invest in their children's human capital. Such underinvestment may lead to higher degrees of sibling resemblance at lower incomes since high ability children from poor families may receive the same low level of education as a sibling with lower academic ability" (cited in Conley and Glauber 2008:300).

2.3.3. Previous Findings

Table 1 summarises current research on within-family inequalities in cognitive abilities and educational outcomes. To my knowledge, only five studies directly assess the association between sibling/twin differences in cognitive ability and educational outcomes in the United States (Grätz and Torche 2016; Griliches 1979), Ethiopia (Ayalew 2005), Mexico (Hussain 2010), and Burkina Faso (Akresh et al. 2012). Only Grätz and Torche (2016) and Hussain (2010) stratify the analyses by parental SES. Griliches (1979) found patterns of slight reinforcement for the effect of a one SD IQ-difference on years of schooling achieved (0.4 to 0.9 years) in the United States. Similarly, Ayalew (2005) in Ethiopia and Akresh and colleagues (2012) in Burkina Faso reported reinforcement trends for children's chances of attending school (probability differences of 0.09 and 16.4, respectively). In Mexico, Hussain (2010) also found (slight) reinforcement for the overall sample in the chances of attending secondary school (probability difference of 0.03). For families with high secondary education, he found a compensating/neutral parental response (–0.02); for non-educated (0.05) and primary-educated (0.02) families, he found a slight reinforcement. Grätz and Torche (2016) found that highly educated parents provide more cognitive stimulation to children with a higher early ability (reinforcement), although this differential response does not explain later cognitive development or school readiness. In turn, early cognitive and motor development at age 10 months has a direct positive impact on cognitive performance at 4 years for low- and high-SES families alike (reinforcement of early ability), net of parental stimulation at 2 years old.

Table 1. Literature review on within-family (in)equalities in cognitive abilities and educational outcomes

Study	Country	IQ → Edu. Outcome	Parental Response	Educational Outcome(s)	Average Result	SES	Parental SES Measure	Design
Griliches 1979	United States	IQ	Indirectly measured	Educational attainment (years of education)	Reinforcement	No	No	Fixed-effects (siblings/twins)
Ayalew 2005	Ethiopia	IQ (Raven test)	Indirectly measured	Attending school	Reinforcement	No	No	Fixed-effects (siblings)
Conley et al. 2007	United States	No	Indirectly measured	Cognitive outcomes (literacy, numeracy, reading comprehension, problem-solving skills): Woodcock-Johnson Revised (WJ-R) Test of Achievement	High SES: <ICC (reinforcement) Low SES: >ICC (compensation)	Yes	-Mother's years of education (<13/>13) -Race (black/white)	ICC (siblings)
Conley & Glauber 2008	United States	No	Indirectly measured	Years of formal schooling completed	High SES: >ICC (compensation) Low SES: <ICC (reinforcement)	Yes	Mother's years of education (<13/>13)	ICC (siblings)
Conley 2008	United States	No	Indirectly measured	Educational attainment	No ICC differences by parental SES	Yes	-Mother's high school degree -Race (black/white)	ICC (siblings)
Hussain 2010 (working paper)	Mexico	IQ (Raven progressive matrices test)	Indirectly measured	-Grade attainment -Grade retention -Age at enrolment -Age quit school -Secondary school	High SES: compensation Low SES: reinforcement	Yes	Parental education	Fixed-effects (siblings)
Akresh et al. 2012	Burkina Faso	IQ (Raven's Colored Progressive Matrices and Wechsler Intelligence Scales Digit Span)	Indirectly measured	-Current enrolment in primary school -Ever enrolled -Grade progression	Reinforcement	No	No	Fixed-effects (siblings)
Grätz & Torche 2016	United States	-Bailey Scales of Infant Development at age 10 months (motor and cognitive) -Early Child Development at age 10 months (crawling, sitting, walking, standing)	Directly measured: Parental cognitive stimulation at 2 years old (Two-Bags test)	-Cognitive performance (maths and reading tests) at 4 years old	Cognitive t0 -> Stimulation t+1 High SES: reinforcement Low SES: neutrality Stimulation t+1 -> Cognitive t+2 High SES: neutrality Low SES: neutrality Cognitive t0 -> Cognitive t+2 High SES: reinforcement Low SES: reinforcement	Yes	-Parental education -Household income -Family SES (high/low)	Fixed-effects (twins)
Grätz 2018	Germany	No	Indirectly measured	-Cognitive performance -Grade point average -Upper-track attendance	No ICC differences by parental SES	Yes	-Parental education and class -Parental ISEI (high/low) -Migration background	ICC (siblings)

The overall picture from previous research points to the reinforcement of cognitive endowments on educational outcomes. Nevertheless, none of these studies looked at the heterogeneity of this association across the absolute ability distribution, and only two stratified analyses by SES.

3. The German Educational System

The design of educational systems may attenuate or reinforce early inequalities generated within families (Landes and Nielsen 2012; Skopek and Passaretta 2018). Educational systems with early-ability tracking reinforce the magnitude of early academic ability on educational inequality (primary effects), whereas comprehensive systems reinforce the role of decisions (secondary effects), as all pupils follow similar tracks during lower-secondary education, and transitions to upper secondary are generally less dependent on previous performance (Blossfeld et al. 2016; Jackson 2013). Early tracking systems seem to lead to larger overall socioeconomic inequalities (Bol and Werfhorst 2013).

The German educational system is decentralised by federal states (länders), but early-ability tracking generally starts at the last grade of joint primary education, at age 10 (grade 4) or 12 (grade 6 at orientation-level schools). At this point, teachers recommend to parents a track choice for their children. Legislation between länders differs greatly regarding the existence or level of binding of the recommendation but recent research shows that the effect of social background on children's chances of accessing the academic track remains fairly stable across different levels of binding recommendations (Roth and Siegert 2016).

After primary grade 4, most pupils have access to three track-specific types of secondary schools: lower-secondary (*hauptschule*), middle-secondary (*realschule*), or upper secondary (*gymnasium*). *Hauptschule* and *realschule* lead to vocational training education, whereas *gymnasiums* offer the most academically-oriented education. The vast majority of *gymnasium* students enter university[48] after passing the *abitur* exam (Schneider 2008). Some states have other types of schools, for example, comprehensive schools (*gesamtschule*) were created in the 1960s by the Social Democratic Party to integrate the three-track system into one school with three tracks. However, in practice, comprehensive schools are considered lower-rank and have not replaced the three-tier system. Even though Germany has relaxed the regional legislation

[48] According to Schneider (2008:512), "Over 90% of the general university entrance qualifications awarded in the year 2001 were obtained by attending Gymnasium."

in tracking age and the binding of track recommendations, and has allowed more horizontal movement between tracks (Blossfeld et al. 2016), the initial tracking allocation is a bottleneck that makes it very difficult to change an individual's educational pathway (Fishkin 2014; Schneider 2008).

Because sizeable academic-ability differences by parental background exist before track sorting (primary effects) (Blossfeld et al. 2017), early-ability tracking fosters "ability or meritocratic selection" (Esser 2016), whereas the comprehensive system leaves more leeway for parental choice. Because teachers are supposed to recommend track allocation on the basis of observed academic performance (i.e., mathematics, German, and classroom behaviour), parents of children with low academic performance or ability may have less room to influence track decisions. Thus, the German case is a stringent test for the compensatory advantage hypothesis, compared to previous research on educational systems without early tracking.

In the event of low-performing children in high-SES families, compensatory patterns could essentially work via parental pressure for a positive recommendation, directly ignoring grades or teachers' positive bias (Schneider 2008). Teachers' recommendations are subject to bias (Boone et al. 2018) by, for instance, misconceiving cultural capital as academic brilliance (Jæger and Møllegaard 2017) or assessing more favourably children with fewer behavioural problems (Møllegaard 2016). Additionally, Jürges and Schneider (2007) claim that low-SES parents are more likely to send their children to vocational tracks even if they have a recommendation for the academic track, whereas the opposite occurs in high-SES families. They argue that this difference may be explained by high-SES parents' higher levels of educational aspirations.

By using cognitive ability as a measure of academic potential, which is less tightly associated with recommendations or track choice than are teacher-assigned grades, compensatory patterns may also work by active parental involvement with the lower-ability twin to improve academic performance (e.g., help with homework and school curriculum, motivation). These are the mechanisms I attempt to isolate by testing the compensatory advantage hypothesis at the bottom of the academic-ability distribution.

4. Data, Variables and Sample Selection

4.1. Data

To answer the aforementioned research questions, I use the first wave of the *Twin Life Study – Genetic and Social Causes of Life Chances*, a cross-sequential panel study comprising four age cohorts of same-sex twins aged 5 (born 2009/10), 11 (born 2003/04), 17 (born 1997/98), and 23-24 (born 1990/93) (Diewald et al., 2018). I study a cohort of same-sex 11-year-old monozygotic (MZ) and dizygotic (DZ) twins at grades 5 and 6 (n = 2,012 twins/1,006 families) born in 2003 and 2004. The first face-to-face wave of the study was carried out between 2014 and 2016, interviewing twins, siblings and parents with CAPI, CASI and PAPI survey methods, with a participation rate of about 40%. The *Twin Life Study* was designed as a probability-based sample intended to be representative of German municipalities and rural areas, and families with same-sex twins (Brix et al., 2017). The sample was drawn from administrative registries of residents by identifying those individuals with identical addresses, birthdays and genders.

Technical reports of the *Twin Life Study* compared distributions of the key socio-demographic variables of the survey with the German micro-census survey by identifying a proxy-twin and a multiple-child household sample (Lang and Kottwitz 2017). The sample of the *Twin Life Study* is positively selected in terms of urban households, German citizenship, parental SES and mothers' age. The distribution of track attendance in the *Twin Life* study is very similar to the that of the general population of students. As can be seen in Table 3 of Chapter IV drawing from NEPS data, about 59% of German students attend the academic track, while in the *Twin Life Study* sample, as shown in Table 3 of this Chapter, this share is slightly lower at 54%. Inequalities by parental education in track attendance are about 10% larger in the *Twin Life* sample even when it is a positively selected sample in terms of SES. In the *Twin Life Study* (NEPS), 73% (70%) of children from highly-educated families attend the academic track, while only 36% (44%) of children from low-educated families do the same. I think that the larger observed inequalities in the *Twin Life* sample may be due to the dissolution of resources in low-SES families in the case of multiple children.

4.2. Variables

Track attendance. The dependent variable on track attendance is measured with a dummy on the type of secondary school currently attended: 0 = *haupshule* and *realshule* (vocational training tracks: comprising lower and intermediate-secondary schools, integrated secondary schools, and comprehensive schools) and 1 = *gymnasium* (academic track: upper secondary schools). I excluded pupils still attending primary education and orientation-level schools, which usually delay tracking decisions until 12 years old. I conducted sensitivity checks by excluding observations of students attending comprehensive schools, and the results are robust.

Cognitive ability. I measure cognitive ability with the Culture Fair Test (CFT 20-R), a widely used and well-validated cognitive test battery that captures non-verbal (fluid) intelligence as a proxy for general cognitive abilities, as the general factor of intelligence also includes verbal ability (Schulz et al. 2017). This test is designed to minimise the influence of sociocultural and environmental factors, although it still reflects them. The test was administered via computer, resulting in a sum of all correctly answered items in a battery of four subtests on figural reasoning (15 items), classification (15 items), matrices (15 items), and reasoning (topology) (11 items). I applied a latent factor approach on the four subtests (Gottschling 2017). The factor analysis (principal components) indicates that the four subtests load strongly on a single component with the following factor loadings by subtest: figural reasoning (0.7424), classification (0.7597), matrices (0.7969), and reasoning (0.5922). I constructed a standardised general cognitive abilities score from these four items with a satisfactory Cronbach's alpha at 0.86.

The test was administered when respondents were in grades 5 or 6. Thus, cognitive abilities were measured at least one grade after tracking (grade 4). Given that education is causally associated with gains in cognitive abilities (Carlsson et al. 2005), I carried out a robustness check (see below) and concluded that overestimation bias is not compromising the results.

Parental background. I measure parental background with a dummy for the highest education level (ISCED-97) achieved by either the father or the mother; 0 = ISCED 1-5B ($\leq$ upper secondary) and 1 = ISCED 5A-6 (university and PhD). I codify this variable in such a reduced way to maximise sample size to split the analysis by parental background, and to compare the children of university graduates vs everyone else. I carried out sensitivity analyses (see Table A.2. in Appendix A.2.) with an alternative measure of parental background, using the highest parental socioeconomic status (ISEI), and the results hold.

Covariates. I control for a set of key variables that may confound the main associations under study, both within families—*z*-birth-weight deviation from pair mean—and between families—twin-pair zygosity (MZ = 0; DZ = 1), twin-pair gender (male = 0; female = 1), *z*-birth-weight pair mean, and mean parental cognitive abilities (as measured by the Culture Fair Test) to approximate the environmental effect of parental education, net of intergenerational genetic transmission of cognitive abilities (Björklund et al. 2010; Conley et al. 2015).

Sample selection. Table 2 describes the cases missing and excluded from the overall sample for this study's variables of interest. The majority of missing cases come from the dependent variable on track attendance (23%). Within the missing cases on track attendance, the majority (47%) were students still attending primary education due to grade retention or orientation-level schools delaying tracking until age 12. The incidence of missing cases on the outcome variable is slightly higher for lower-medium-educated families (+6.3%), due in part to a larger prevalence of grade retention. Cognitive abilities and BW each account for 8.9% of missing cases. After applying list-wise deletion[49] on the variables of interest, 36.6% of cases are excluded, with a larger incidence for lower-medium-educated families (40%) than for highly-educated ones (34%). In Appendix A.1., I discuss in more detail the possible sample selection bias and conclude that inequalities may be underestimated due to the positive selection of the analytic sample.

[49] I did not apply multiple imputation to solve the problem of small sample size, missing data, and attenuation bias in twin fixed-effects because there are not good auxiliary variables that vary within families predicting missing information in the outcome.

Table 2. Sample selection

Variables	All Sample		Low-Med. Edu.		High-Edu.	
	n missing	% missing	*n* missing	% missing	*n* missing	% missing
Academic track	478	22.96%	282	25.54%	192	19.28%
Zygosity	0	0%	0	0%	0	0%
Gender	0	0%	0	0%	0	0%
z-Birth weight*	186	8.93%	102	9.24%	82	8.23%
z-Cognitive abilities*	186	8.93%	108	9.78%	78	7.83%
z-Parental cognitive abilities	50	2.40%	34	3.08%	16	1.61%
Highest parental education	12	0.58%				
Samples	*n*	%	*n*	%	*n*	%
Overall sample*	2,082	100%	1,104	100%	996	100%
Excluded cases*	762	36.60%	442	40.04%	338	33.94%
Analytic sample	1,320	63.40%	662	59.96%	658	66.06%

*Including non-missing cases within unbalanced twin-pairs.

4.3. Empirical Strategy

4.3.1. Identification Strategy: Twins as a Natural Experiment

An ideal test of the compensatory advantage hypothesis would compare siblings who differ in nothing but their (observable) academic potential.[50] Nature provides an experimental setting with the incidence of twins (Knopik et al. 2017). Twins are a quasi-random phenomenon, being born into the same family on the same day and sharing at least 50% of their genetic makeup. Hence, twin fixed-effects models rule out most sources of variation within families that might confound the association between cognitive abilities and educational outcomes. Twin-pairs discordant in exposure can be thought of as a natural counterfactual in which the co-twins can be used as their own control/experimental group (McGue et al. 2010).

This design allows for the control of more unobserved confounding than most previous research that uses between-family estimates and sibling fixed-effects (Jæger 2011) but within-family variation in cognitive abilities might not be randomly assigned. Three potential sources of variation might confound the association between cognitive ability and track choice within families: pre-natal, genetic (DZ-twins), and unique environmental (DZ and MZ-twins) (Knopik et al. 2017).

[50] Measures of parents' perceptions of twins' cognitive abilities show that parents can fairly well identify which twin has the greater or fewer cognitive abilities.

First, to account for prenatal environmental factors (placenta position and type, access to oxygen and nutrients *in utero*) that may confound the association between twin differences in cognitive abilities and track allocation, the analyses control for BW as a proxy for endowments at birth. Prior research shows that BW is associated with early cognitive development and educational attainment (Almond and Mazumder 2013).

Second, genetic differences between DZ-twins can influence twin differences in cognitive abilities and track allocation (Conley et al. 2015). Genetic sources of variation in cognitive abilities within a family are a random phenomenon, given that, in the process of reproduction, each sibling randomly gets 50% of their segregating alleles from each parent (Knopik et al. 2017). However, one could argue that genetic sources of variation in personality traits (e.g., attention control) associated with both cognitive abilities and educational outcomes may confound the association under study. The causal links and sources of co-variation between the development of non-cognitive and cognitive abilities are currently far from clear (Meldrum et al. 2017). Additional twin fixed-effects models controlling for twins' (and parents') concentration and persistence abilities show the main effects of cognitive abilities are robust and of similar effect size as this non-cognitive measure.

Third, unique or non-shared environmental factors may affect the development of cognitive abilities and track choice differently for each co-twin (Asbury, Moran and Plomin 2016) (e.g., twin-specific parental preferences, investment, or stimulation, or twin-specific reactions to peer effects, friends, or teachers). The bias would be positive if twin-specific factors affect cognitive abilities and track choice in the same direction: bias would be negative if twin-specific factors affect cognitive abilities and track choice in the opposite direction. Regarding twin-specific parental response, to my knowledge, the only study to look at within-family associations between early cognitive abilities, parental responses, and later cognitive abilities did not find a direct association between the last two factors among a cohort of twins in the United States (Grätz and Torche 2016). My analyses suggest that intra-class correlations of cognitive ability, and the distribution of cognitive-ability differences within families, for twins from low- and high-SES families do not differ systematically (see Table 3). Finally, concerning the effects of twin-specific friends, classmates, and teachers, sensitivity analyses show that placing twins in different classrooms (family-constant dummy affecting 43% of the sample) during primary school neither mediates nor moderates the association between twin differences in cognitive abilities and twin differences in track choice.

4.3.2. Hybrid Multilevel Models: Between-Within Estimators

Given the paired structure of the data, I implement multilevel models comprising two twins (level-1) clustered in families (level-2), only keeping balanced pairs at level-1. As I am interested in estimating and comparing both between- and within-family parameters, I use hybrid multilevel models (also known as between-within models) (Carlin et al. 2005; McGue et al. 2010; Turkheimer and Harden 2014). These models include the twin-pair average ($\beta_2(\overline{X_j}) = \frac{X_{i1j} + X_{i2j}}{2}$), and the deviation from the twin-pair average ($\boldsymbol{\beta_1}(X_{ij} - \overline{X_j})$), for each twin for the variables that vary within families (BW and cognitive abilities). The former parameter is largely equivalent to a naïve or pooled OLS regression in which twins are treated as individual observations; the latter is equivalent to a standard twin fixed-effects model that just accounts for variation within families, controlling for family-constant factors. Finally, because cross-level interactions are estimated for the deviation of cognitive abilities and the pair-mean cognitive abilities, all models include random slopes for the level-1 variable interacted (deviation of cognitive abilities) (Heisig and Schaeffer 2018). I estimate linear probability models (LPM) in all specifications for the sake of comparability and interpretation.

$$y_{ij} = \alpha + \beta_1(X_{ij} - \overline{X_j}) + \beta_2(\overline{X_j}) + \beta_3(\boldsymbol{Z}_{ij} + \boldsymbol{Z}_j) + \mu_j + e_{ij} + e(X_{ij} - \overline{X_j}) \quad (1)$$

Equation 1 shows the baseline hybrid multilevel model in which y_{ij} measures type of track attendance for twin *i* in family *j*; α represents the intercept or grand mean probability of accessing the academic track across families; β_1 stands for the main coefficient of interest on the (fixed) effect of twin *ij*'s cognitive ability deviation from the pair-mean cognitive ability in family *j* within discordant twin-pairs; β_2 stands for the effect of pair-mean cognitive abilities in family *j*; and β_3 represents a vector of covariates between and within families. On the right-hand side of the equation, the random-effects parameters represent the error term decomposed in a between- and within-family component: μ_j is the pair-level error of prediction, or unaccounted variance between families; e_{ij} stands for the individual/within-family error of prediction, or the individual unaccounted share of the variance; and $e(X_{ij} - \overline{X_j})$ expresses the random slopes' parameter for the cognitive ability deviation. This parameter represents to what extent the effect of cognitive-abilities deviation varies across families.

I estimate four different models. Model 1, as expressed in equation 1, estimates the effect of within- and between-family differences in cognitive abilities and parental background on the probability of attending the academic track. To answer research question 1, I estimate the model expressed in equation 1 independently for lower-medium-educated (Model 2a) and highly-educated (Model 2b) families.

To answer research question 2, equation 2 includes β_4, which represents the cross-level interaction between cognitive-abilities deviation from pair-mean and pair-mean cognitive abilities. Therefore, to test the compensatory advantage hypothesis within families, we must assess whether the fixed-effect of twin differences in cognitive abilities on track choice (β_1) is heterogeneous across the cognitive-ability distribution (β_2), being *compensated* ($\beta_1 \leq 0$) at the bottom. I estimate two sub-specifications of equation 2 independently for lower-medium-educated (Model 3a) and highly-educated (Model 3b) families.

$$y_{ij} = \alpha + \beta_1(X_{ij} - \overline{X_j}) + \beta_2\,(\overline{X_j}) + \beta_3(\boldsymbol{Z}_{ij} + \boldsymbol{Z}_j) + \beta_4\left(\overline{X_j} * (X_{ij} - \overline{X_j})\right) + \mu_j + e_{ij} + e(X_{ij} - \overline{X_j}) \quad (2)$$

All models are estimated with maximum likelihood, unstructured covariance to allow the within- and between-family residual variance to be correlated: random slopes for β_1 to allow its effect to vary across families, and robust standard errors to better account for non-normally distributed residuals at the family level. Finally, to estimate the share of variance between families, I rely on the intra-class correlation (ρ), as expressed in equation 3 in which $\sigma^2_{\mu_j}$ accounts for between-family variance and $\sigma^2_{e_{ij}}$ for within-family variance.

$$\rho = \frac{\sigma^2_{\mu_j}}{\sigma^2_{\mu_j} + \sigma^2_{e_{ij}}} = \text{Intra-class correlation (ICC)} \quad (3)$$

4.3.3. Within-Family (In)equality: Compensation and Reinforcement

As shown in Table 1, previous literature has tested within-family (in)equalities directly and indirectly. Following the former approach, one can measure early endowments at t_0, parental response at t_{+1}, and educational outcomes at t_{+2}. This empirical strategy is ideal but rare, given that high-quality panel data with rich information on endowments and parental behaviour is needed from birth to the first important educational crossroad (Grätz and Torche 2016).

Regarding the latter approach, most previous research measures within-family (in)equality indirectly in two ways. First, some investigations evaluate the degree of siblings' resemblance

(ICC) in a given socioeconomic outcome, and its potential stratification by parental SES (Conley 2008). Second, other studies assess the effect of sibling differences in a given endowment on later educational outcomes (Ayalew 2005). I apply this second indirect strategy by defining two intra-family patterns in the slope of twins' fixed-effect cognitive abilities on track attendance (β_1): equality or compensation if $\beta_1 \leq 0$, and inequality or reinforcement if $\beta_1 > 0$.

5. Descriptive Analysis

Table 3 shows the descriptive statistics for all variables of the analyses, stratified by parental education. On average, 54% of pupils attend the academic track[51]. There are large differences due to parental education. Just 36% of students from low-medium-educated families follow an academic track, compared to 73% of students from highly-educated families. Previous research on full-siblings shows that within-family variance accounts for 50% of total variance in track placement in Germany (Grätz 2018); I find an estimate of 12.3% (see Table 3).[52] In contrast to some previous findings that siblings from high-SES families have greater resemblance in educational attainment (Conley 2008), the share of total educational attainment variance explained within families does not vary considerably by parental background. Table 3 shows an average ICC in cognitive abilities of 0.64, with virtually no variation by parental background. Figure 1 illustrates the distribution of (within-family) absolute twin differences in cognitive abilities (left-hand side) and deviations from pair-average cognitive abilities (right-hand side). This is the main source of within-family variation that I utilise in this study. On average, absolute twin differences in cognitive abilities stand at 0.77, with a SD of 0.6, with slight variation by SES (see Table 3).

[51] The distribution of track attendance in the *Twin Life Study* is very similar to the one of the general population of students. As can be seen in Table 3 of Chapter IV drawing from NEPS data, about 59% of German students attend the academic track, while in the *Twin Life Study* sample, as shown in Table 3, this share is slightly lower at 54%. Inequalities by parental education in track attendance are about 10% larger in the *Twin Life* sample even when it is a positively selected sample in terms of SES. In the *Twin Life study (*NEPS), 73% (70%) of children from highly-educated families attend the academic track, while only 36% (44%) of children from low-educated families do the same. I would hypothesise that the larger inequalities observed in *Twin Life* may have something to do with the dissolution of resources in low-SES families in case of multiple children.

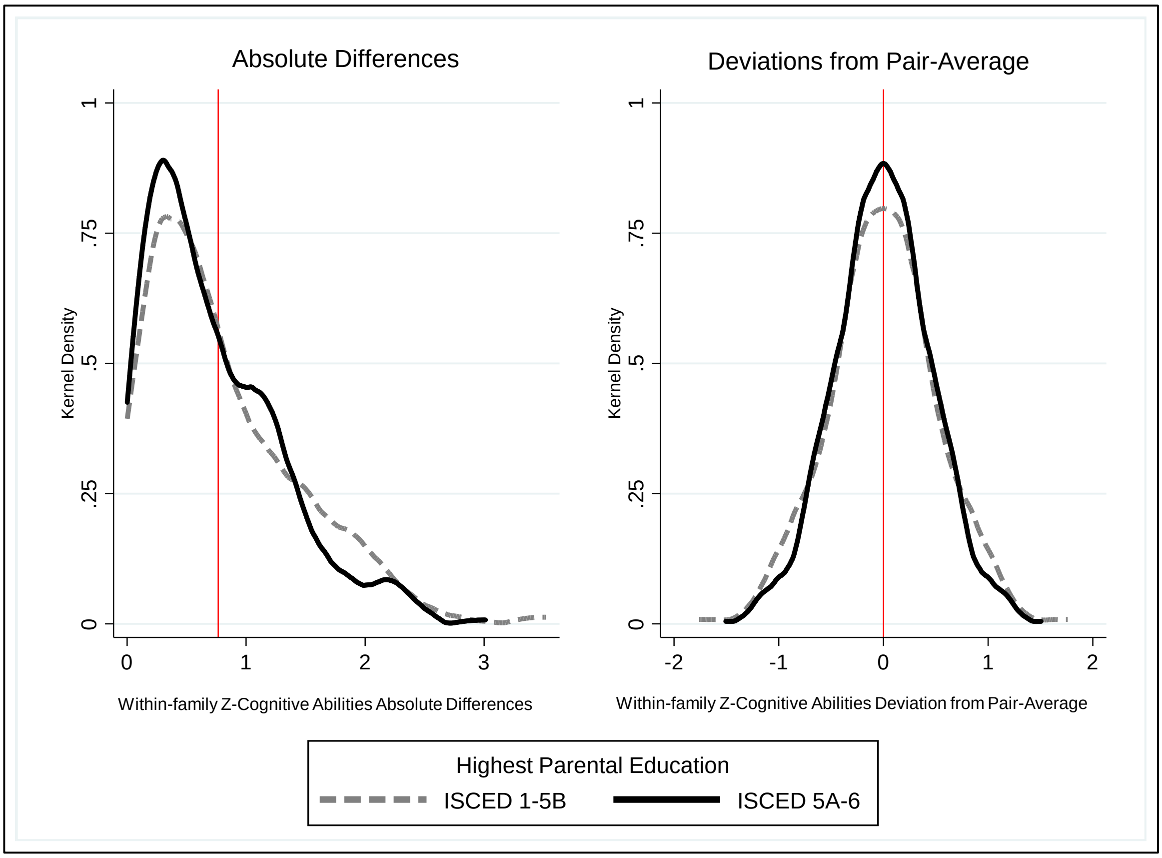
[52] This substantial difference may be due to the fact that (1) the twin-design partially rules out several sources of variation within-families or (2) external validity problems, so that parents or teachers may treat twins more equally (Tully et al. 2004).

Table 3. Analytic sample summary statistics

Variables	Mean	Std. Dev.	Min.	Max.	ICC[a]	Mean	Std. Dev.	Min.	Max.	ICC[a]	Mean	Std. Dev.	Min.	Max.	ICC[a]
	All Sample					Low-Med. Parental Education (ISCED 1-5B)					High Parental Education (ISCED 5A-6)				
Academic track	0.542		0	1	0.877	0.358		0	1	0.870	0.728		0	1	0.835
Dizygotic twin-pairs	0.603		0	1		0.556		0	1		0.650		0	1	
Female twin-pairs	0.521		0	1		0.547		0	1		0.495		0	1	
z-Birth weight	0.058	0.982	-3.143	3.487	0.874	-0.001	1.017	-3.050	3.339	0.893	0.119	0.941	-3.143	3.487	0.850
z-Birth weight pair-mean	0.058	0.925	-3.119	3.311		-0.001	0.967	-3.022	3.311		0.119	0.877	-3.119	2.191	
z-Birth weight deviation	0.000	0.329	-1.667	1.667		0.000	0.317	-1.218	1.218		0.000	0.340	-1.667	1.667	
z-Birth weight abs. diff.	0.481	0.449	0.000	3.333		0.475	0.420	0.000	2.435		0.486	0.476	0.000	3.333	
z-Cognitive abilities	0.115	0.946	-3.508	2.576	0.639	-0.061	0.988	-3.508	2.576	0.628	0.291	0.867	-2.854	2.327	0.613
z-Cognitive-abilities pair-mean	0.115	0.811	-2.520	2.321		-0.061	0.844	-2.520	2.321		0.291	0.736	-1.568	1.995	
z-Cognitive-abilities deviation	0.000	0.488	-1.760	1.760		0.000	0.515	-1.760	1.760		0.000	0.459	-1.505	1.505	
z-Cognitive-abilities abs. diff.	0.766	0.604	0.001	3.519		0.805	0.642	0.001	3.519		0.726	0.560	0.001	3.011	
z-Parental cognitive abilities	0.071	0.911	-3.420	2.033		-0.291	0.969	-3.420	1.575		0.434	0.676	-2.752	2.033	
High parental education	0.498		0	1											
n Twins	1,320					662					658				
n Families	660					331					329				

[a]ICC = intra-class correlation from one-way random effects. Coefficients statistically significant at p<0.001. Notes: abs. diff. = absolute differences.

Figure 1. Distribution of twin differences and deviations from pair-average in cognitive abilities



6. Findings

Table 4 summarises the main findings of this research. Model 1 sheds light on between- and within-family dynamics. Net of child and parental differences in cognitive abilities, twins from highly-educated backgrounds are 27% more likely to attend the academic track compared to their least advantaged counterparts. The effect of parental education (coefficient of 0.36 before controls for cognitive ability) on track allocation is mainly exerted (74.2%) net of parents' and children's cognitive abilities. This suggests that other unobserved factors, net of cognitive abilities, that vary between families with different socioeconomic resources (e.g., risk aversion to downward mobility, non-cognitive abilities) and mediate the association between parental education and track attendance (e.g., via grades, teachers' recommendations, or bias), may account for these observed inequalities. Cognitive abilities account for only around 14% of the variance in track allocation, so there is ample room for other factors that could explain early educational success.

In contrast to the compensatory advantage mechanism found in between-family models, the differences in transition rates by parental education remain constant across the (pair-mean) cognitive-ability distribution. The interaction term between parental education and average cognitive abilities across families is neither substantial nor statistically significant, as formally tested in panel 4 of Table 4 and illustrated in Figure 2. However, the main contribution of this research is to complement this result by testing the compensatory advantage hypothesis within families with different socioeconomic resources.

Table 4-A. Hybrid multilevel LPM with maximum likelihood, random slopes, and unstructured covariance

Outcome: Academic Track	Model 1	Model 2a	Model 2b	Model 3a	Model 3b
Independent Variables	All Sample	Low-Med. Education	High Education	Low-Med. Education	High Education
z-cognitive-abilities deviation	0.0625***	0.0622**	0.0625**	0.0756***	0.0719**
	(0.0153)	(0.0208)	(0.0227)	(0.0221)	(0.0237)
z-cognitive-abilities pair-mean	0.175***	0.168***	0.175***	0.168***	0.178***
	(0.0214)	(0.0300)	(0.0311)	(0.0300)	(0.0312)
Parental *z*-cognitive abilities	0.0454*	0.0635*	0.0149	0.0639*	0.0147
	(0.0207)	(0.0257)	(0.0381)	(0.0258)	(0.0381)
z-cognitive-abilities deviation* *z*-cognitive-abilities pair-mean				0.0643*	-0.0618*
				(0.0253)	(0.0259)
High parental education	0.268***				
	(0.0361)				
Constant	0.331***	0.353***	0.592***	0.354***	0.591***
	(0.0353)	(0.0442)	(0.0476)	(0.0443)	(0.0476)
ICC	.7882	.8199	.7572	.8185	.7549
R^2 level-1 Snijders/Bosker	0.2501	0.1485	0.1136	0.1510	0.1157
R^2 level-2 Snijders/Bosker	0.2763	0.1630	0.1257	0.1630	0.1257
R^2 level-1 Bryk/Raudenbush^	0.0378	0.0369	0.0399	0.0592	0.0550
R^2 level-2 Bryk/Raudenbush^	0.3099	0.1818	0.1427	0.1785	0.1397
AIC	1039.153	539.386	516.173	535.049	514.560
Observations	1,320	662	658	662	658
Number of families	660	331	329	331	329

Note: Controls for twin-pair gender, zygosity, and BW (pair-mean and deviation); robust standard errors are in parentheses.
^Variance explained compared to the null model.
*p<0.05; **p<0.01; ***p<0 .001

Table 4-B. Random-effects parameters

Random-Effects Parameters	Model 1		Model 2a		Model 2b		Model 3a		Model 3b	
	Var.	RSE	Var.	RSE	Var.	RSE	Var.	RSE	Var.	RSE
Var(within family)	0.0379	0.0071	0.0337	0.0108	0.0340	0.0105	0.0411	0.0096	0.0415	0.0096
Var(between family)	0.1410	0.0081	0.1534	0.0111	0.1533	0.0110	0.1280	0.0119	0.1278	0.0119
Var(*z*-cog.-ability dev.)	0.0316	0.0121	0.0352	0.0209	0.0317	0.0191	0.0304	0.0157	0.0272	0.0145
Cov(z-cog.-abilities dev. between family)	-0.0037	0.0034	0.0005	0.0047	0.0028	0.0047	-0.0083	0.0054	-0.0080	0.0056

Note: Var. = variance; Cov. = covariance; RSE = robust standard errors.

Table 4-C. Interaction effects and Wald tests

Wald test for interactions	Coeff. (RSE)	χ^2 (df)	p-value
z-cognitive-abilities dev. × parental edu.	.0004 (.0306)	0.00 (1)	0.990
z-cognitive-abilities pair-mean × parental edu.	-.0116 (.0402)	0.08 (1)	0.772
z-cognitive-abilities pair-mean × z-cognitive-abilities dev. × parental edu.	-.1234** (.0362)	11.65 (1)	0.001

Note: RSE = robust standard errors.
*p<0.05; **p<0.01; ***p<0.001

Model 1 in Table 4 also shows that twin differences in cognitive abilities are predictive of track attendance differences within families (fixed effect at 6.3 percentage points),[53] although this effect is of less magnitude than between families (marginal effect at 17.5 percentage points). More substantially, this coefficient also tells us that intra-family differences in cognitive abilities tend to produce intra-family inequalities in educational attainment. This result is in line with most previous findings that report reinforcement patterns (see Table 1).

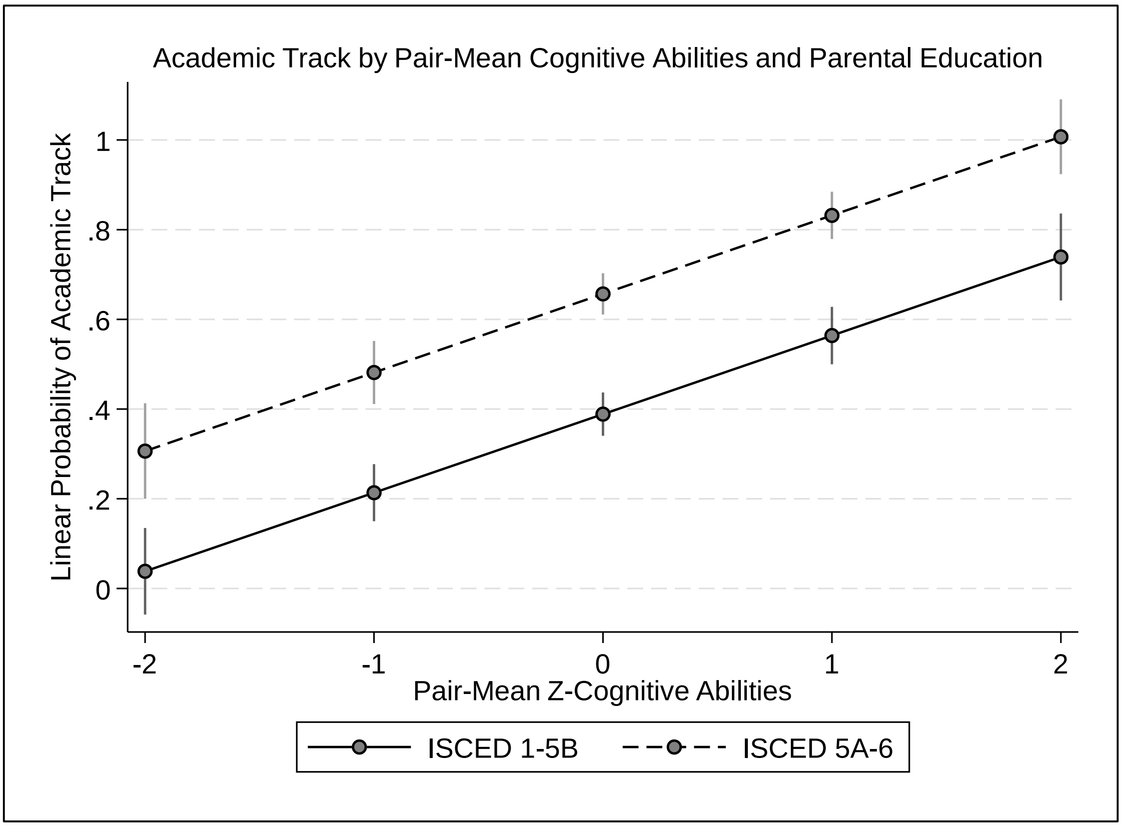


Figure 2. Predicted probabilities at observed values of academic track attendance by pair-mean cognitive abilities and parental education with 95% C.I. (see Table 4, Model 1)

[53] As explained in Appendix A.3., the within-family coefficient is subject to attenuation; its true value should be around 9 percentage points.

Regarding research question 1 on the potential heterogeneity of within-family dynamics by parental background, I find that the effect of twin differences in cognitive abilities on track placement is not stratified by parental education at average absolute levels of cognitive abilities (see Table 4, Models 2a and 2b; formally tested in panel C of Table 4 with an interaction). Contrary to some previous theoretical predictions and findings (Conley and Glauber 2008; Hussain 2010), results suggest that twin differences in cognitive abilities tend to produce within-family inequalities in educational outcomes among advantaged and disadvantaged families alike:[54] twins with greater cognitive abilities show larger transition rates (+6 %) to the academic track than do their co-twins with lesser academic potential.

The main drawback of previous research theories and findings is that within-family associations by parental SES may be contingent on children's absolute level of endowments. Thus, the compensatory advantage hypothesis should be tested at the bottom of the absolute academic-ability distribution. To do so, Models 3a and 3b in Table 4 display the cross-level interaction between twin differences in cognitive abilities (deviation) and pair-average cognitive abilities (absolute distribution). In Models 3a (low-medium-educated parents) and 3b (highly-educated parents), this interaction term is statistically significant and of similar substantial magnitude but of different sign (0.06 and –0.06, respectively). The difference between both interaction coefficients by SES is statistically significant, as formally tested in Table 4, panel C. This result means that within-family (in)equalities depend on twin-pairs' absolute level of cognitive abilities.

Figure 3 displays the predicted probabilities at observed values for this interaction term. Overall, this figure shows a more fine-grained picture than do previous theories and findings. In both low-medium- and highly-educated families, twins' differences in cognitive ability generate within-family equality (*compensation*) and inequality (*reinforcement*) patterns. As Figure 4 shows in the fixed-effects slopes across the (twin-pair) cognitive-ability distribution, in disadvantaged families, twin differences in cognitive abilities lead to within-family equality at the bottom of the cognitive-ability distribution ($\boldsymbol{\beta} = 0.00$), but they produce inequality at the middle ($\boldsymbol{\beta} = 0.07$) and, especially, at the top ($\boldsymbol{\beta} = 0.14$). Advantaged families show the opposite pattern. Twin differences in cognitive ability generate the largest within-family

[54] The random parameters in Table 4, panel B, illustrate that the residual variance between and within families is of similar magnitude by parental education. Thus, systematic sources of unobserved heterogeneity that would compromise the observed patterns can be ruled out (i.e., non-random measurement error in ability).

inequalities at the bottom of the cognitive-ability distribution (***β*** = 0.14), more modest inequalities at the middle (***β*** = 0.07), and equality at the top (***β*** = 0.00).

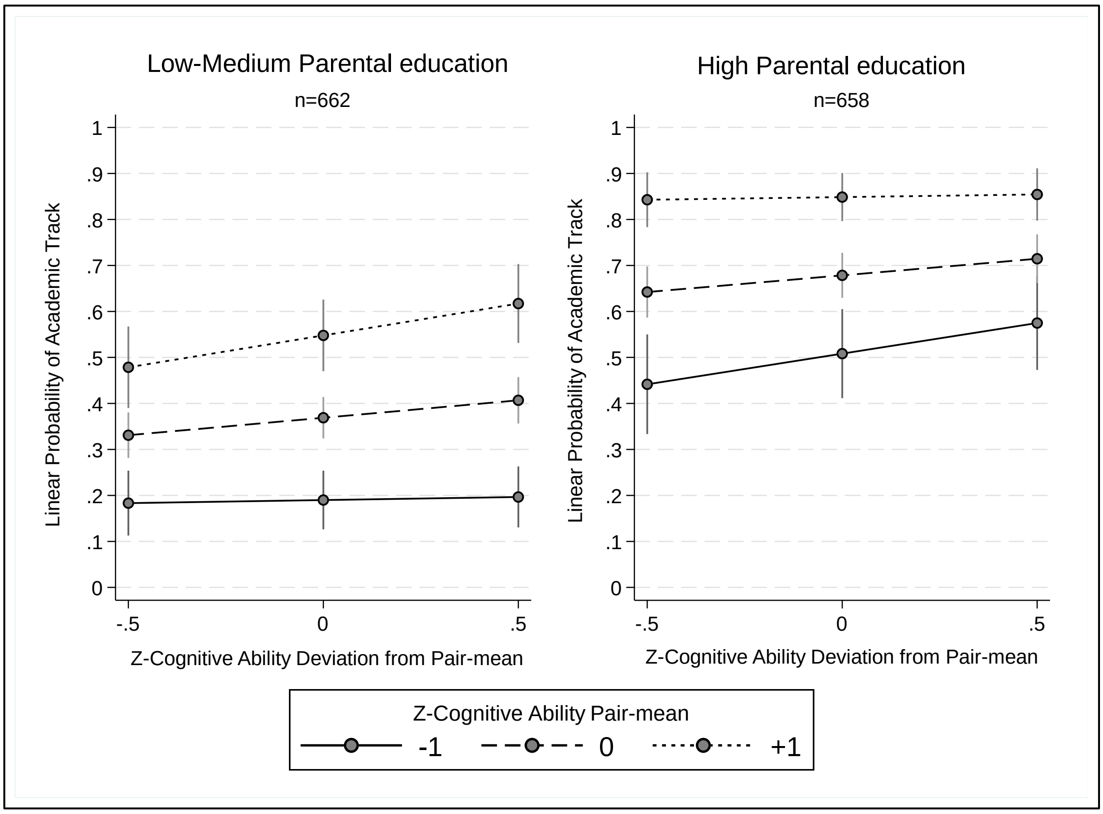


Figure 3. Predicted probabilities at observed values and random effects = 0 for the interaction between *z*-cognitive-abilities deviation (fixed-effect) and pair-mean cognitive abilities by parental education (see Table 4, Models 3a and 3b) with 95% C.I.

These intra-family patterns, across the absolute cognitive-ability distribution, point to the compensatory advantage mechanism going in the opposite direction in the German educational system. Namely, it seems highly-educated families are not able to compensate for children's low academic ability, as lower-ability twins at the bottom of the cognitive-ability distribution show the largest differences in transition rates with respect to their relatively more gifted co-twins. One might think the absence of compensatory patterns in advantaged families is good news for equality of opportunity. Nonetheless, as we saw in Figure 2, children from highly-educated families still have substantially larger transition rates to the academic track.

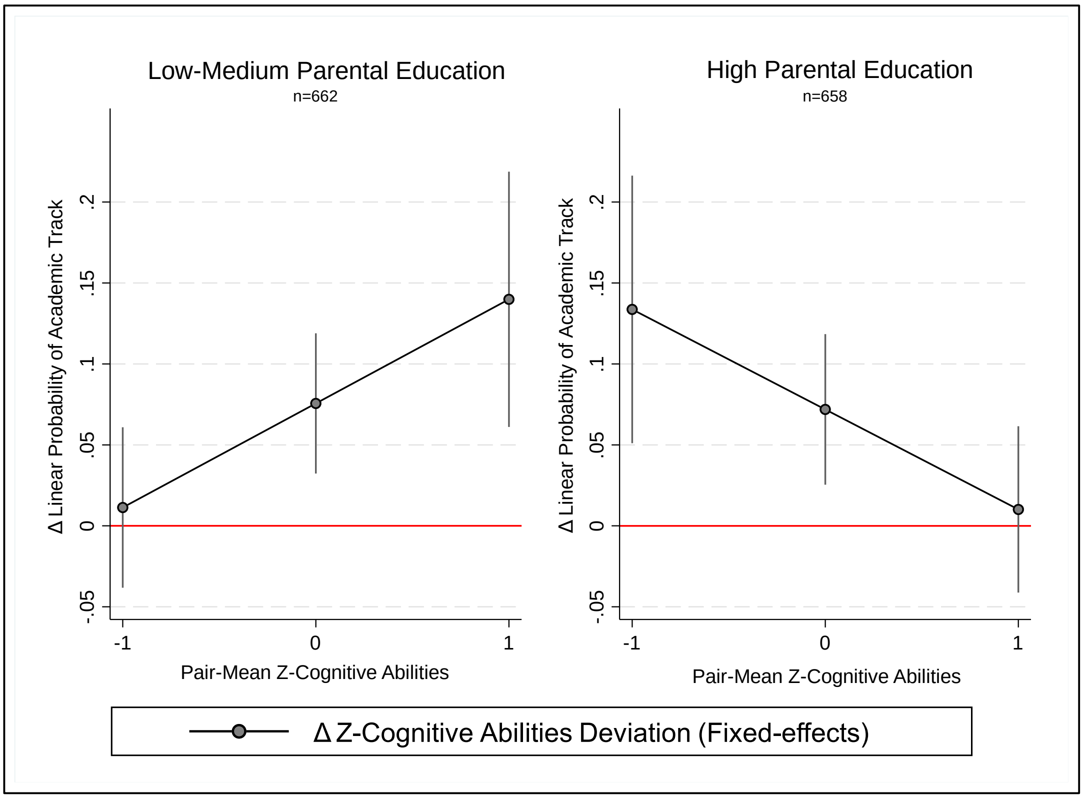


Figure 4. Interaction between *z*-cognitive-abilities deviation (AME or fixed effects) and pair-mean cognitive abilities (absolute cognitive-ability distribution, moderator in *X*-axis) by parental education (see Table 4, Models 3a and 3b) with 95% C.I.

Rational action theories have mainly been developed and applied to studying educational inequalities between families, but I argue that the theorised mechanisms that differ between low-, medium-, and highly-educated families (resources and risk-aversion to downward mobility) may help us to understand these opposite patterns of within-family (in)equalities. In the German educational system, and others like it, in which the recommendation or transition to secondary education is mainly dependent on early academic ability, highly-educated parents may have difficulty compensating for twins' ability differences at the bottom of the ability distribution. The risk of downward mobility might be at a maximum at this threshold; hence, ability differences may be magnified so that at least the higher-ability twin makes it to the academic track.

In contrast, for lower- and medium-educated families at the bottom of the academic-ability distribution, it does not matter how much relative ability a twin may display, because both twins have equally low chances of attending the academic track (*equality to the bottom*). However, at higher levels of the ability distribution, twin differences in cognitive abilities may become more noticeable for parents, thus generating inequality patterns as parents attempt to help the higher-ability twin make it to the academic track. The main logic behind this speculation is that disadvantaged families are generally more reluctant to opt for the academic track given their lower resources, especially in the case of families with twins, and lower perceived chances of success. Unless one of their children is exceptionally bright in absolute and relative terms, the parents are more risk-averse.

7. Robustness Checks

7.1. Reverse Causality

Cognitive ability is measured at least one grade (5th or 6th) after tracking (4th grade) takes place. Previous research shows a positive longitudinal association between academic tracking and gains in cognitive ability (Guill, Lüdtke and Olaf 2017). Even though the longitudinal intra-personal correlation of cognitive abilities is very high in the short-term ($r \approx .80$), and its main sources of stability over time are genetic in origin (Briley and Tucker-Drob 2013; Deary and Johnson 2010), an association between academic tracking and gains in cognitive ability could compromise accurate estimations of the effect of cognitive ability on track choice due to potential reverse causality bias (i.e., overestimation). More importantly, for this research, in the case of reverse causality, reinforcement patterns could be more easily found than compensation.

Reverse causality would ideally be tested by assessing differences in cognitive ability in the academic and vocational tracks before (grade 4) and after (grade 5 onward) tracking (at 10 years old). Unfortunately, I cannot observe cognitive abilities before tracking. Thus, to estimate the direction and magnitude of this potential reverse causality bias, I utilise a feature of German national legislation for enrolment in primary school: children must turn 6 on or before June 30 to enrol in school, although there is variability by länders (Jürges and Schneider 2007). This cut-off based on birth-month generates variation in grade progression in the sample. Those kids who turn 6 before the cut-off are enrolled in the first grade of primary, while those who turn later on delay their enrolment until the next academic year. Even though allocation to grade is not completely based on random variation coming from pupils' birth-months, as families have a considerable margin of discretion (Bernardi 2014), I use this variation in grade progression to assess reverse causality.

The main idea is to assess whether pupils in the academic track increase their advantage in cognitive ability compared to students in vocational tracks between grades 5 and 6. After excluding observations that experienced grade retention and twins attending different grades, I compare average cognitive-ability differences by track between grades 5 and 6 with naïve OLS (equation 4) and fixed-effects (FE) (equation 5) regressions. Cognitive ability and the dummies on grade (X_i) and track (Z_i) are interacted, showing whether the difference in cognitive ability between academic and vocational tracks increased, decreased, or remained constant between grades 5 and 6.

$$y_i = \alpha + \beta_1(X_i) + \beta_2(Z_i) + \beta_3(X_i * Z_i) + e_i \quad (4)$$

$$y_{i1j} - y_{i2j} = \beta_1(x_{i1j} - x_{i2j}) + \beta_2[(x_{i1j} - x_{i2j}) * (z_j)] + (e_{i1j} - e_{i2j}) \quad (5)$$

As Table 5 and Figure 5 show, pupils in vocational training and academic tracks increased their mean cognitive ability from grades 5 to 6. However, while the advantage in cognitive ability for academic-track pupils at grade 5 stands at 0.67 (naïve-OLS) and 0.56 SD-units (FE), this advantage does not significantly increase by grade 6. The magnitude of the difference in cognitive abilities between tracks remains fairly stable across grades 5 and 6, as shown in the coefficients for the interaction terms in Table 5: –0.11 SD-units for the OLS estimator, and –0.06 SD-units for the fixed-effect estimator. These differences are not statistically significant.

Table 5. Reverse causality: effects of track and grade on cognitive ability

Outcome: Cognitive Ability	Naïve OLS	Twin-FE
Grade 6 (Grade 5)	0.3357***	
	(0.0893)	
Academic track (VT)	0.6662***	0.5641**
	(0.0918)	(0.2047)
Track × grade	-0.1070	-0.0625
	(0.1157)	(0.2652)
Constant	-0.1070***	-0.1349
	(0.0702)	(0.0784)
Observations	1,029	1,029
R^2	0.1261	0.0322

Note: Reference categories and standard errors are in parentheses; robust SE for the naïve OLS model. *p<0.05; **p<0.01; ***p<0.001

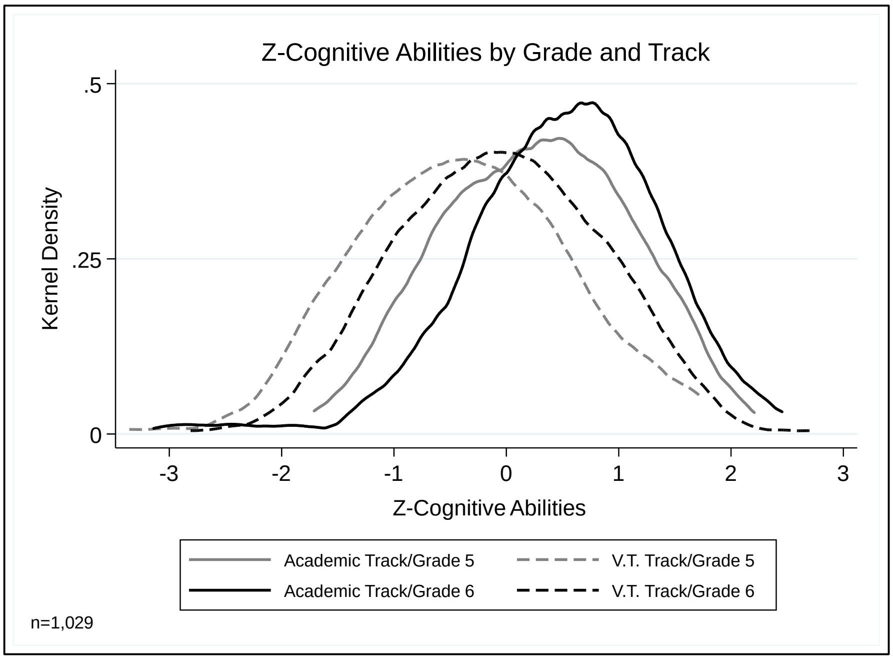


Figure 5. Cognitive-ability distribution by grade and track

Even if this robustness check cannot completely rule out reverse causality problems (i.e., unobserved cognitive abilities between grades 4 and 5), it shows that overestimation bias might not represent a serious threat for this article's general finding: children with low cognitive abilities from high-SES families experience no compensation.

7.2. Attenuation and Measurement Error

An additional concern about twin fixed-effects estimations revolves around measurement error in the main independent or exposure variable of interest: cognitive ability. In this paper, cognitive ability is measured by *The Culture Fair test* (CFT 20-R), which is a sum of all correctly answered items from a battery of four subtests on figural reasoning, classification, matrices, and topology. This test is a proxy for an underlying latent concept of general cognitive abilities, with a satisfactory reliability Cronbach's alpha at 0.86. Previous literature has shown that IQ tests have a measurement error of around 10% (Sandewall et al. 2014).

Under classical random measurement error theory, the signal-to-noise ratio in within-family or fixed-effects estimations is much lower than in between-family or pooled OLS estimations. As most variation in cognitive ability is produced between-families, within-family estimates are more subject to measurement error bias towards zero, especially so among MZ-twins. As Kohler et al. (2012:14) point out, "if the coefficient estimate from the fixed-effects twins estimator is smaller, it may be because it controls for the endogenously determined part of cognitive ability, or because of the larger bias due to measurement error or due to some combination of these two factors." For even with a large enough sample size with adequate power to detect statistically significant effects, within-pair estimates are attenuated due to the compounding of the error, making it more difficult to find within-pair associations even when there is a causal effect.

Since I have neither valid instruments for cognitive abilities in the survey, nor independent estimations of the level of measurement error in cognitive ability for the *Culture Fair Test*, we cannot directly estimate the signal-to-noise ratio or the reliability ratio of cognitive ability in the sample analysed. Nonetheless, drawing from own analyses, between-family estimates (0.175) of the association between cognitive abilities and track allocation are 64% larger than within-family estimates (0.063), and the average twin-correlation in cognitive ability (*g*) stands at 0.64 $(corr(g_1, g_2))$. Relying on these figures and the noise-to-signal ratios theorised by Kohler et al. (2012:15), it can be assumed a noise-to-signal ratio of around 0.16, or, alternatively, a cross-sectional reliability ratio (*r*) of *g* of around 0.84. Assuming classical measurement errors, the within-pair reliability ratio or the attenuation factor can be imputed by applying the following equation (Sandewall et al., 2014:5).

$$\textit{Imputed within-pair reliability ratio} = \frac{(r - corr(g_1,g_2))}{(1 - r\ corr(g_1,g_2))} \quad (6)$$

By substituting this formula by the assumed cross-sectional reliability ratio (*r*) of *g* at 0.84 and the observed twin-correlation in *g* at 0.64, the imputed within-pair reliability ratio equals $\frac{0.84 - 0.64}{1 - 0.64} = 0.44$. This means that within-pair estimations of cognitive abilities on track allocation are downwardly biased by a 0.44 factor, approximately. Thus, it should be noted that the main within-family coefficients of interest represent a lower-bound threshold and should be tentatively corrected by an additional 0.44 of its value. For instance, the association between twin-differences in cognitive abilities and twin-differences in track allocation stand at 0.063; thus, the correction for the attenuation effect would equal 0.063*0.44=0.028. Hence, the corrected coefficient of interest would equal 0.063+0.028=0.091. By implication and assuming no unobserved heterogeneity, the true causal parameter of the effect of cognitive abilities on track allocation would lie between 0.091 (corrected within-family coefficient) and 0.20 (corrected between-family coefficient; 0.175*0.16= 0.028; 0.175+0.028=0.20).

7.3. Additional Robustness Checks

To further assess the credibility of the findings, in Appendix A.2. I discuss additional robustness checks on external validity, logistic specifications, extrapolation and linearity of the moderation analysis, and alternative measures of parental SES.

8. Conclusions

The main aim of this article was to test whether high-SES families compensate for children's low ability in the transition to secondary education in the stringent setting of German early-ability tracking. This article was motivated by the lack of dialogue and limitations of the literature on educational inequalities between (i.e., misspecification of social background and ability) and within (i.e., stratification by SES and the endowment distribution) families. I used a twin fixed-effects design that controls for more unobserved confounding (i.e., school, genes, and neighbourhood) and provides a more credible test of the compensatory hypothesis than most previous research using between-family estimates or sibling-FE.

I find that twins with greater cognitive abilities than their co-twins enjoy larger transition rates to the academic track. This finding aligns with previous research that finds reinforcement patterns for this association in the United States, Mexico, Ethiopia, and Burkina Faso. Does parental SES moderate the effect of twin differences in cognitive ability on track choice (research question 1)? The positive association between cognitive ability and transition to the academic track, generating within-family inequality or reinforcement of abilities, holds for advantaged and disadvantaged families alike. In other words, in contrast to some previous hypotheses and findings (Conley and Glauber 2008), within-family inequality in educational outcomes is not heterogeneous across parental SES. This result is in line with Grätz's (2018) study in Germany that finds no SES-heterogeneity in the level of siblings' similarity in admission to the academic track.

The main aim of this article was to test the compensatory advantage hypothesis within families at the bottom of the absolute cognitive-ability distribution (research question 2). Results show that highly-educated families are not able to compensate for children's low academic ability: lower-ability twins at the bottom of the ability distribution show the largest differences in transition rates compared to their relatively more able twin. Rational action theories (i.e., risk aversion to downward mobility), normally applied to understand between-family inequalities, may also help us understand these within-family patterns.

In the German educational tracking system, in which the recommendation and transition to secondary education is thought to be a function of early academic ability, highly-educated parents may have difficulties deploying compensatory strategies at the bottom of the academic-ability distribution. It remains a question whether patterns of compensation for low cognitive abilities may emerge in educational systems without early-ability tracking, in which transitions to upper secondary are less linked to observed performance, or when using more direct measures of academic ability such as competencies or GPA.

The absence of compensatory patterns in cases of low academic ability might be interpreted as positive evidence for equality of opportunity, but children from highly-educated families with the same level of cognitive ability as children from less-educated ones still have substantially larger transition rates to the academic track. From a normative standpoint, these inequalities net of cognitive ability represent a waste of academic potential for disadvantaged students, which compromises upward social mobility. Moreover, this scenario is at odds with the role of cognitive ability as a prominent criterion of merit for liberal theories of equality of opportunity, especially so in the German system of ability-tracking (Fishkin 2014).

Overall, results point to the importance of other unobserved factors, rather than cognitive ability, in influencing learning, academic performance, and transition rates that vary between families with different socioeconomic resources. Potential factors that could be responsible for this residual association between parental background and educational outcomes include risk-aversion to downward mobility, non-cognitive skills, and teachers' bias. Further research is needed to explore these mechanisms.

After carrying out several robustness checks on reverse causality, attenuation and sample selection bias, moderation, confounding, and alternative specifications, I generally conclude that the study's main findings are consistent. A substantive limitation of this study is that no direct indicators of parental investment or responses to children's endowments are used. Future research should disentangle the particular mechanisms that may account for the associations between children's endowments, parental response, and educational outcomes across families with different socioeconomic resources. Recognising that these limitations should be improved in future research, this article contributes to the literature on educational inequality between and within families in theoretical, methodological, and empirical terms.

APPENDIX

A.1. Sample Selection Bias

Table A.1. displays a mean-comparison for the overall and analytic samples stratified by parental education to assess the potential selection bias of the analytic sample. There are small differences between the average values and standard deviations of the overall and analytic samples, but the samples are generally representative. Note, however, that lower-medium-educated families decrease their share in the total sample by 3%, and, in the analytic sample, their average cognitive abilities with respect to the overall sample increase by around 18% for children and by 10% for parents. This means that, if anything, the main analyses slightly underestimate socioeconomic inequalities.

Table A.1. Means and mean-differences between overall and analytic samples for the full sample and by parental background

Variables	Full Sample			Low-Medium Education			High Education		
	Analytic	Overall	Diff.	Analytic	Overall	Diff.	Analytic	Overall	Diff.
Academic track	0.542	0.520	2.21%	0.358	0.341	1.74%	0.728	0.721	0.70%
Dizygotic twin-pair	0.603	0.596	0.74%	0.556	0.565	-0.93%	0.650	0.629	2.11%
Female twin-pair	0.521	0.520	0.15%	0.547	0.549	-0.21%	0.495	0.487	0.89%
z-Birth weight	0.058	0.000	5.84%	-0.001	-0.077	7.53%	0.119	0.088	3.09%
z-Birth weight pair-mean	0.058	-0.005	6.38%	-0.001	-0.080	7.84%	0.119	0.080	3.83%
z-Birth weight deviation	0.000	0.000	0.00%	0.000	0.000	0.00%	0.000	0.000	0.00%
z-Birth weight abs. diff.	0.481	0.500	-1.94%	0.475	0.494	-1.91%	0.486	0.505	-1.84%
z-Cognitive abilities	0.115	0.000	11.47%	-0.061	-0.233	17.22%	0.291	0.277	1.44%
z-Cognitive abilities pair-mean	0.115	-0.018	13.25%	-0.061	-0.259	19.82%	0.291	0.271	2.00%
z-Cognitive abilities deviation	0.000	0.000	0.00%	0.000	0.000	0.00%	0.000	0.000	0.00%
z-Cognitive abilities abs. diff.	0.766	0.787	-2.14%	0.805	0.816	-1.06%	0.726	0.756	-3.02%
z-Parental cognitive abilities	0.071	0.005	6.52%	-0.291	-0.388	9.66%	0.434	0.460	-2.57%
High parental education	0.498	0.467	3.18%						

Note: abs. diff = absolute differences

A.2. Additional Robustness Checks

Due to space limitations, I cannot show all the detailed analyses (available from the author upon request) but can only summarise the following key bullet points. (1) Even when parents and teachers may be averse to treating twins differently, and twin fixed-effects only account for a small portion of the total variance in cognitive abilities and track choice in comparison to full-sibling models, I find reinforcement patterns that were much more difficult to find in the case of external validity problems. (2) Logistic specifications yield highly equivalent results in magnitude to the linear probability models, favouring the use of the latter to maximise comparability between models, power, and sample size. The interpretation of the coefficients as marginal effects or changes in probabilities is much straightforward than odds ratios (Albertini et al. 2018:4-5). Furthermore, the odds ratios estimated in hybrid multilevel models are cluster-specific, yielding very high odds for the cluster-variables that are complex to interpret. I carried out the same hybrid multilevel models of the main analysis using logistic specifications, and the direction, magnitude (interaction term's odds ratio=4.1 for lower-medium-educated parents, and 0.23 for highly-educated parents) and statistical significance (for the cross-level interaction effects, the *p-values* range from 0.010 for lower-medium-educated parents, to 0.057 for highly-educated-parents) of the coefficients of interest is relatively equivalent to the LPM specifications. Finally, an analysis of the distribution of the predicted probabilities after running LPM specifications shows that just 2.6% of the predictions are out-of-range in the LPM, and the baseline of the outcome (54%) makes it analogous to the logistic function. (3) As shown in Figures A.1. and A.2., using linear, categorical and kernel specifications, the cross-level interaction effects meet reasonably well the assumption on linearity across the estimated range (−1/+1 SD) of the moderator (pair-mean cognitive abilities), and the sample is large enough at these points to extrapolate findings (Hainmueller, Mummolo and Xu 2018). Thus, to maximise sample size, a linear specification is preferred instead of breaking down the moderator into tertiles. (4) Analysis using an alternative measure of parental SES (i.e., ISEI) yields highly equivalent results to parental education, as shown in Table A.2.

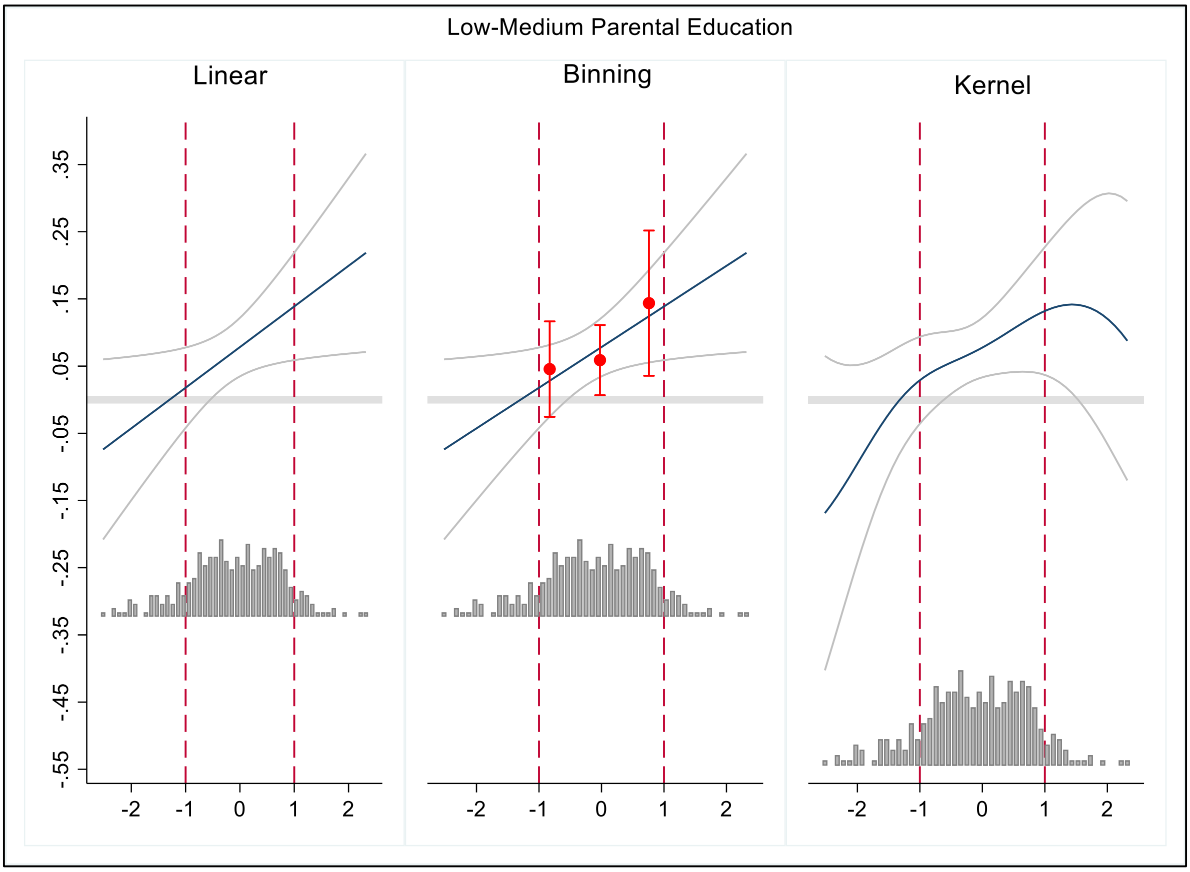


Figure A.1. Interaction between *z*-cognitive-abilities deviation (AME or fixed effects in the *Y*-axis) and pair-mean cognitive abilities (absolute cognitive-ability distribution, moderator in *X*-axis) for low-medium educated families with 95% C.I. using linear, binning and kernel specifications of cognitive ability. Controls and sample size as in Models 3a and 3b.

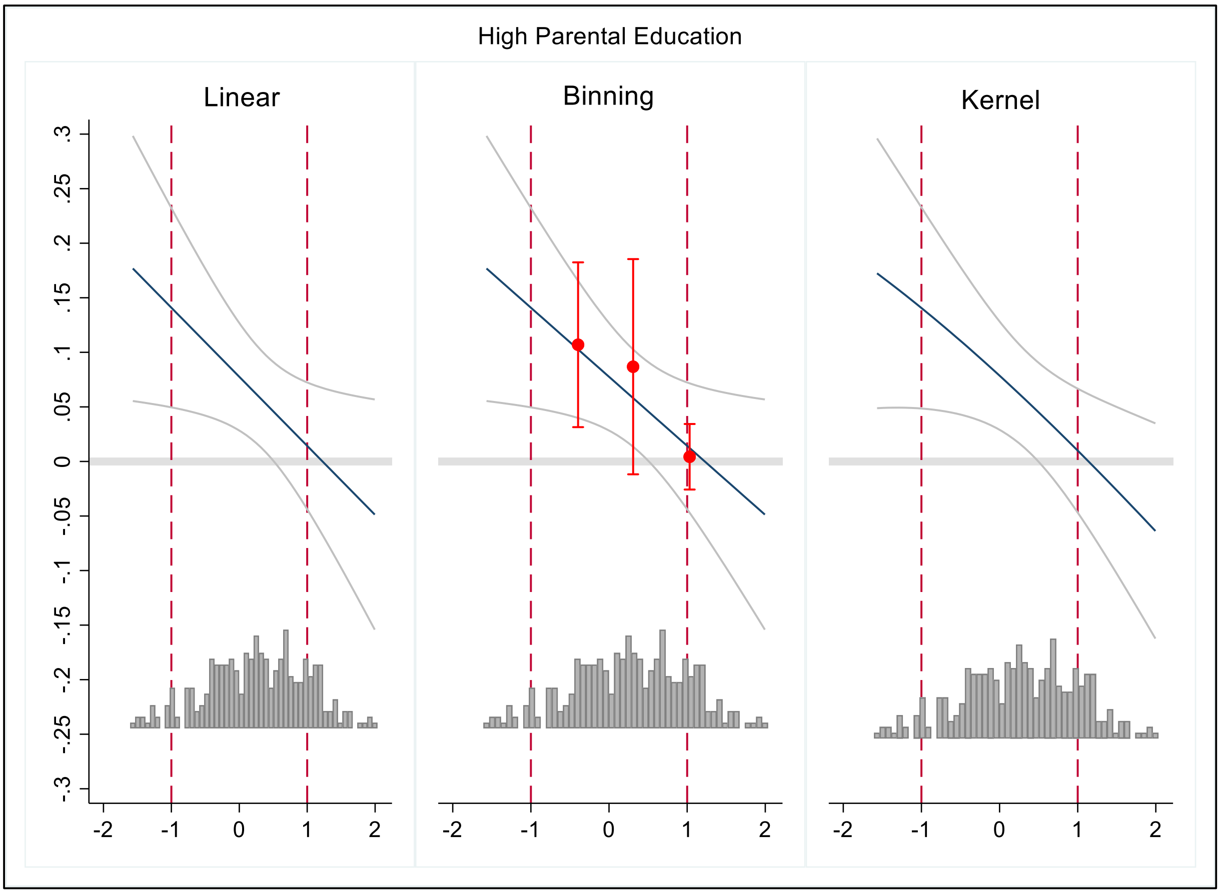


Figure A.2. Interaction between z-cognitive-abilities deviation (AME or fixed effects in the Y-axis) and pair-mean cognitive abilities (absolute cognitive-ability distribution, moderator in X-axis) for highly-educated families with 95% C.I. using linear, binning and kernel specifications of cognitive ability. Controls and sample size as in Models 3a and 3b.

Table A.2. Hybrid multilevel LPM with maximum likelihood, random slopes, and unstructured covariance

Outcome: Academic Track Independent Variables	Model 1 All Sample	Model 2a Low ISEI	Model 3b High ISEI	Model 3a Low ISEI	Model 3b High ISEI
z-cognitive-abilities deviation	0.0556***	0.0462*	0.0632**	0.0593**	0.0776**
	(0.0154)	(0.0203)	(0.0229)	(0.0207)	(0.0251)
z-cognitive-abilities pair-mean	0.165***	0.149***	0.181***	0.148***	0.184***
	(0.0220)	(0.0317)	(0.0301)	(0.0318)	(0.0301)
Parental *z*-cognitive abilities	0.0657*	0.0668*	0.0706+	0.0670**	0.0701
	(0.0230)	(0.0292)	(0.0389)	(0.0293)	(0.0389)
z-cognitive-abilities deviation X *z*-cognitive-abilities pair-mean				0.0580*	-0.0660**
				(0.0245)	(0.0248)
High parental education	0.198***				
	(0.0367)				
Constant	0.372***	0.386***	0.546***	0.387***	0.546***
	(0.0375)	(0.0486)	(0.0474)	(0.0486)	(0.0474)
ICC	0.7868	0.7931	0.7757	0.7929	0.7749
R^2 level-1 Snijders/Bosker	0.2127	0.116	0.15	0.1183	0.1524
R^2 level-2 Snijders/Bosker	0.2352	0.1274	0.1667	0.1274	0.1667
R^2 level-1 Bryk/Raudenbush^	0.0304	0.0214	0.0447	0.0428	0.0624
R^2 level-2 Bryk/Raudenbush^	0.2639	0.142	0.1898	0.1391	0.1864
AIC	1027.487	561.9636	478.162	558.292	475.949
Observations	1,258	634	624	634	624
Number of families	629	317	312	317	312

Note: Controls for twin-pair gender, zygosity, and BW (pair-mean and deviation); robust standard errors are in parentheses.
^Variance explained compared to the null model.
*p<0.05; **p<0.01; ***p<0.001

References

Akresh, Richard, Emilie Bagby, Damien de Walque, and Harounan Kazianga. 2012. "Child Ability and Household Human Capital Investment Decisions in Burkina Faso." *Economic Development and Cultural Change*, 61(1):157–186.

Albertini M, Gähler M, and Härkönen J. 2018. Moving back to "mamma"? Divorce, intergenerational coresidence, and latent family solidarity in Sweden. *Popul Space Place.*

Almlund, Mathilde, Angela L. Duckworth, James J. Heckman, and Tim D. Kautz. 2011. "Personality psychology and economics." Pp. 1-181 in *Handbook of the Economics of Education*, vol. 4, edited by E.A. Hanushek, S. Machin, and L. Wößmann, L. Amsterdam: Elsevier.

Almond, Douglas, and Bhashkar Mazumder. 2013. "Fetal Origins and Parental Responses." *Annual Review of Economics*, 5:37-56.

Asbury, Kathryn, Nicola Moran, and Robert Plomin. 2016. "Nonshared Environmental Influences on Academic Achievement at Age 16: A Qualitative Hypothesis-Generating Monozygotic-Twin Differences Study." *AERA Open*, 2(4).

Ayalew, Tekabe. 2005. "Parental preference, heterogeneity, and human capital inequality. *Economic Development and Cultural Change.*" 53(2):381-407.

Bailey, Drew, Greg J. Duncan, Candice L. Odgers, and Winnie Yu. 2017. "Persistence and Fadeout in the Impacts of Child and Adolescent Interventions." *Journal of Research on Educational Effectiveness*, 10(1):7–39.

Becker, Gary S. and Nigel Tomes. 1976. "Child endowments and the quantity and quality of children." *Journal of Political Economy*, 84(4):S143-S162.

Becker, Gary S. and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3):1–39.

Behrman, Jere R., Robert A. Pollak, and Paul Taubman 1982. "Parental Preferences and Provisions for Progeny." *Journal of Political Economy*, 90(1):52.

Bernardi, Fabrizio. 2014. "Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth." *Sociology of Education*, 87(2):74–88.

Bernardi, Fabrizio, and Héctor Cebolla-Boado. 2014. "Previous School Results and Social Background: Compensation and Imperfect Information in Educational Transitions." *European Sociological Review*, 30(2):207-214.

Bernardi, Fabrizio, and Moris Triventi. 2018. "Compensatory advantage in educational transitions: Trivial or substantial? A simulated scenario analysis." *Acta Sociologica*, doi:10.1177/0001699318780950

Björklund, A., Eriksson, K. H., & Jäntti, M. 2010. IQ and family background: Are associations strong or weak? *The BE Journal of Economic Analysis & Policy*, 10(1), 1–12.

Blossfeld, Hans-Peter, Sandra Buchholz, Jan Skopek, and Moris Triventi, eds. 2016. *Models of secondary education and social inequality. An international compariso*n. Cheltenham (UK) and Northampton (MA, USA): Edward Elgar.

Blossfeld, Hans-Peter, Nevena Kulik, Jan Skopek, and Moris Triventi, eds. 2017. *Childcare, Early Education and Social Inequality. An International Perspective.* Cheltenham (UK) and Northampton (MA, USA): Edward Elgar.

Bol, Thijs, and Herman G. van de Werfhorst. 2013. "Educational Systems and the Trade-Off between Labor Market Allocation and Equality of Educational Opportunity." *Comparative Education Review*, 57(2):285–308.

Boone, Simon, Sarah Thys, Piet Van Avermaet, and Mieke Van Houtte. 2018. "Class composition as a frame of reference for teachers? The influence of class context on teacher recommendations." *British Educational Research Journal*, 44(2):274–293.

Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*, New York: Wiley.

Bowles, Samuel and Herbert Gintis. 2002. "Schooling in Capitalist America Revisited." *Sociology of Education*, 75:1-18

Breen, Richard and John H. Goldthorpe. 1997. "Explaining Educational Differentials: towards a Formal Rational Action Theory." *Rationality and Society*, 9(3):275- 305.

Briley. D. A.. and Tucker-Drob, E. M. 2013. Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, *24*, 1704-1713.

Brix, J., Pupeter, M., Rysina, A., Steinacker, G., Schneekloth, U., Baier, T., Gottschling, J., [...] & Spinath, F. M. (2017). A longitudinal twin family study of the life course and individual development (TwinLife): Data collection and instruments of wave 1 face-to-face interviews. TwinLife Technical Report Series, 05. Project TwinLife: Genetic and social causes of life chances (Universität Bielefeld / Universität des Saarlandes).

Bukodi, Erzsébet, Robert Erikson, and John H. Goldthorpe. 2014. "The effects of social origins and cognitive ability on educational attainment: Evidence from Britain and Sweden." *Acta Sociologica*, 57(4):293–310.

Bukodi, Erzsébet, Mollie Bourne, and Bastian Betthäuser. 2017. "Wastage of talent?: Social origins, cognitive ability and educational attainment in Britain." *Advances in Life Course Research*, 34:34–42.

Carlin, John B., Lyle C. Gurrin, Jonathan A.C. Sterne, Ruth Morley, and Terry Dwyer. 2005. "Regression models for twin studies: a critical review." *International Journal of Epidemiology*, 34:1089-1099.

Carlsson, Magnus, Gordon B. Dahl, Björn Öckert, and Dan-Olof Rooth. 2015. "The effect of schooling on cognitive skills." *Review of Economics and Statistics*. 97(3):533-547.

Conley, Dalton. 2004. *The Pecking Order: Which Siblings Succeed and Why*. New York: Pantheon.

Conley, Dalton. 2008. "Bringing Sibling Differences In: Enlarging Our Understanding of the Transmission of Advantage in Families." Pp. 179-200, in *Social Class: How Does It Work?*, edited by D. Conley and A. Lareau A. New York: Russell Sage Foundation.

Conley, Dalton, and Rebecca Glauber. 2008. "All in the family?: Family composition, resources, and sibling similarity in socioeconomic status." *Research in Social Stratification and Mobility*, 26(4):297–306.

Conley, Dalton, Kathryn M. Pfeiffer, Melissa Velez, M. 2007. "Explaining sibling differences in achievement and behavioral outcomes: The importance of within- and between-family factors." *Social Science Research*, 36(3):1087–1104.

Conley, Dalton, Benjamin W. Domingue, David Cesarini, Christopher Dawes, Cornelius A. Rietveld, and Jason D. Boardman.2015. "Is the Effect of Parental Education on Offspring Biased or Moderated by Genotype?" *Sociological Science*, 2:82-105.

Deary, Iian J., and Wendy Johnson. 2010. "Intelligence and education: causal perceptions drive analytic processes and therefore conclusions." *International Journal of Epidemiology*, 39(5):1362-1369.

Deary, Ian J., Steve Strand, Pauline Smith, and Cres Fernandes. 2007. "Intelligence and educational achievement." *Intelligence*, 35:13–21.

Diewald, Martin, et al. 2018. *TwinLife. GESIS Data Archive, Cologne. ZA6701 Data file Version 2.1.0*, doi:10.4232/1.13072.

Duckworth, Angela L., Patrick D. Quinn, and Eli Tsukayama. 2012. “What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card.” *Journal of Educational Psychology*, 104:439–451.

Erikson, Robert, and Frida Rudolphi. 2010. “Change in Social Selection to Upper Secondary School—Primary and Secondary Effects in Sweden.” *European Sociological Review*, 26(3):291–305.

Erola, Jani, and Elina Kilpi-Jakonen, eds. 2017. *Social Inequality across the Generations. The Role of Compensation and Multiplication in Resource Accumulation*. Cheltenham, UK: Edward Elgar Publishing.

Esping-Andersen, Gösta, and Jorge Cimentada. 2018. “Ability and mobility: The relative influence of skills and social origin on social mobility.” *Social Science Research*, 75:13–31.

Esser, H. 2016. “The model of ability tracking – Theoretical expectations and empirical findings on how educational systems impact on educational success and inequality.” Pp. 25-42, in *Models of Secondary Education and Social Inequality. An International Comparison*, edited by H.-P. Blossfeld, S. Buchholz, J. Skopek, and M. Triventi. Cheltenham (UK) and Northampton (MA, USA): Edward Elgar.

Euro-Peristat Project (2008). *European Perinatal Health Report. Core indicators of the health and care of pregnant women and babies in Europe in 2004*. www.europeristat.com

Farah, Marta J., et al. 2008. “Environmental stimulation, parental nurturance and cognitive development in humans.” *Developmental Science*, 11(5):793-801.

Fishkin, Joseph. 2014. *Bottlenecks: A New Theory of Equal Opportunity*. Oxford: Oxford University Press.

Goldthorpe, John H. 2007. *On Sociology*. Stanford, California: Stanford University Press.

Gottschling, Juliana. 2017. *Documentation TwinLife Data: Cognitive Abilities* (TwinLife Technical Report Series, 02). Bielefeld: Project TwinLife "Genetic and social causes of life chances" (Universitat Bielefeld).

Gottschling, Juliana, Elisabeth Hahn, Christopher Beam, Frank M. Spinath, Sarah Carroll, and Erik Turkheimer. 2019. Socioeconomic status amplifies genetic effects in middle childhood in a large German twin sample. *Intelligence*, 72:20–27.

Grätz, Michael. 2018. “Competition in the Family: Inequality between Siblings and the Intergenerational Transmission of Educational Advantage.” *Sociological Science*, 5:246-269.

Grätz, Michael, and Florencia Torche. 2016. "Compensation or Reinforcement? The Stratification of Parental Responses to Children's Early Ability." *Demography*, 53:1883-1904.

Griliches, Zvi. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *The Journal of Political Economy*, 87(5):S37-S64.

Guill, Karin, Oliver Lüdtke, and Olaf Köller. 2017. "Academic tracking is related to gains in students' intelligence over four years : Evidence from a propensity score matching study." *Learning and Instruction*, 47:43-52.

Guo, Guang, and Elizabeth Stearns. 2002. "The Social Influences on the Realization of Genetic Potential for Intellectual Development." *Social Forces*, 80(3): 881-910.

Hahn, Elisabeth, et al. 2016. "What Drives the Development of Social Inequality Over the Life Course? The German TwinLife Study." *Twin Research and Human Genetics*, 19(6):659-672.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2018. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*. https://doi.org/10.1017/pan.2018.46

Härkönen, Juho. 2014. "Birth Order Effects on Educational Attainment and Educational Transitions in West Germany." *European Sociological Review*, 30(2):166–179.

Heisig, Jan P., and Merlin Schaeffer. 2018. "Why You Should Always Include a Random Slope for the Lower-level Variable Involved in a Cross-level Interaction." *SocArXiv*. January 11. doi:10.31235/osf.io/bwqtd.

Herrnstein, R. J., and Murray, C. (1994). The Bell Curve: Intelligence and Class Structure in American Life. NY.

Hussain, Iftikhar. 2010. "Parental preferences and inequality within the family: evidence from Mexican siblings." *Working Paper. University of Oxford: Oxford.* Retrieved from https://core.ac.uk/download/pdf/9836416.pdf

Jackson, Michelle, eds. 2013. *Determined to succeed? Performance versus choice in educational attainment*, Stanford: Stanford University Press.

Jæger, Mads M. 2011. "Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data." *Sociology of Education*, 84(4):281–298.

Jæger, Mads M., and Stine Møllegaard. 2017. "Cultural capital, teacher bias, and educational success: New evidence from monozygotic twins." *Social Science Research*, 65:130–144.

Jensen, A. R. 1969. How much can we boost IQ and scholastic achievement? Harvard Educational Review. 39, 1–123. doi:10.17763/haer.39.1.l3u15956627424k7

Jürges, Hendrik, and Kerstin Schneider, 2007. “What Can Go Wrong Will Go Wrong: birthday Effects and Early Tracking in the German School System.” *CESIfo Working Paper 2055.* CESifo. Munich. Retrieved from http://www.cesifo-group.de/DocDL/cesifo1_wp2055.pdf

Kendler, Kenneth S., Eric Turkheimer, Henrik Ohlsson, Jan Sundquist, and Kristina Sundquist. 2015. “Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study.” *Proceedings of the National Academy of Sciences*, 112(15):4612–4617.

Knopik, Valerie S., Jenae N. Neiderhiser, John C. DeFries, and Robert Plomin. 2017. *Behavioral genetics.* New York: Macmillan Learning.

Kohler, Hans-Peter, Jere R. Behrman. and Jason Schnittker. 2011. “Social science methods for twins data: Integrating causality, endowments, and heritability.” *Biodemography and Social Biology*, 57(1):88–141.

Landes, Xavier, and Morten E.J. Nielsen. 2012. “Intra-Family Inequality and Justice.” *Dialogue*, 51(3):437-466.

Lang, Volker, and Anita Kottwitz. 2017. *The sampling design and socio-demographic structure of the first wave of the TwinLife panel study: a comparison with the Microcensus* (TwinLife Technical Report Series, 03) updated version, August 2017. Bielefeld: Project TwinLife "Genetic and social causes of life chances" (Universität Bielefeld / Universität des Saarlandes).

Lynch, Jamie L., and Ryan Brooks. 2013. Low birth weight and parental investment: Do parents favor the fittest child? *Journal of Marriage and Family*, 75(3):533–543.

McGue, Matt, Merete Osler, and Kaare Christensen. 2010. “Causal Inference and Observational Research: The Utility of Twins.” *Perspectives on Psychological Science*, 5(5):546–556

Meldrum, Ryan C., Melissa A. Petkovsek, Brian B. Boutwell, and Jacob T.N. Young. 2017. “Reassessing the relationship between general intelligence and self-control in childhood.” *Intelligence*, 60(Supplement C):1–9.

Møllegaard, Stine. 2016. “The Role of Cultural Capital and Behavioral Problems in Educational Inequality.” Ph.D. Dissertation. Faculty of Social Sciences, University of Copenhagen, Copenhagen.

Neisser, Ulric, et al. 1996. “Intelligence: Knowns and Unknowns.” *American Psychologist*, 51:77-101.

Nisbett, Richard E., Joshua Aronson, Clancy Blair, William Dickens, James Flynn, Diane F. Halpern, and Eric Turkheimer. 2012. "Intelligence: new findings and theoretical developments." *The American Psychologist*, 67(2):130-159.

Papageorge, Nicholas W., and KevinThom. 2018. "Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study." *NBER Working Papers* 25114, National Bureau of Economic Research, Inc.

Poropat, Arthur E. 2009. "A meta-analysis of the five-factor model of personality and academic performance." *Psychological Bulletin*, 135:322–338.

Ritchie, Stuart J., and Elliot M. Tucker-Drob. 2018. "How Much Does Education Improve Intelligence? A Meta-Analysis." *Psychological Science*, 29(8):1358–1369.

Rindermann, Heiner. 2007. "The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS and IQ-Tests across Nations." *European Journal of Personality*, 21(5):667–706.

Roth, Tobias, and Manuel Siegert. 2016. "Does the Selectivity of an Educational System Affect Social Inequality in Educational Attainment? Empirical Findings for the Transition from Primary to Secondary Level in Germany." *European Sociological Review*, 32(6):779–791.

Sandewall, Örjan, David Cesarini, and Magnus Johannesson. 2014. "The co-twin methodology and the returns to schooling – testing a critical assumption." *Labour Economics*, 26:1–10.

Schneider, Thorsten. 2008. "Social Inequality in Educational Participation in the German School System in a Longitudinal Perspective: Pathways into and out of the most Prestigious School Track." *European Sociological Review*, *24*(4):511–526.

Schulz, Wiebke, Reinhard Schunck, Martin Diewald and Wendy Johnson. 2017. "Pathways of Intergenerational Transmission of Advantages during Adolescence: Social Background, Cognitive Ability, and Educational Attainment." *Journal of Youth and Adolescence*, 46(10):2194–2214.

Sewell, William H., and Vimal P. Shah. 1968. "Parents' Education and Children's Educational Aspirations and Achievements." *American Sociological Review*, 33(2):191-209.

Sieben, Inge, and Paul de Graaf. 2003. "The total impact of the family on educational attainment." *European Societies*, 5(1):33–68.

Skopek, Jan, and Giampiero Passaretta. 2018. "The Social Stratification of Skills from Infancy to Adolescence – Evidence from an Accelerated Longitudinal Design." *SocArXiv*. October 1. doi:10.31235/osf.io/xkctv.

Stocké, Volker 2007. "Explaining Educational Decision and Effects of Families' Social Class Position: An Empirical Test of the Breen–Goldthorpe Model of Educational Attainment." *European Sociological Review*, *23*(4):505–519.

Tucker-Drob, Elliot M., Daniel A. Briley, and K. Paige Harden. 2013. "Genetic and environmental influences on cognition across development and context." *Current Directions in Psychological Science*, 22:349-355.

Tully Lucy, Terrie E. Moffitt, Avshalom Caspi, Alan Taylor, Helena Kiernan, and Penny Andreou. 2004. "What effect does classroom separation have on twins' behavior, progress at school, and reading abilities?" *Twin Research*, 7(2):115-124.

Turkheimer, Erik, and K. Paige Harden. 2014. "Behavior genetic research methods: Testing quasi-causal hypotheses using multivariate twin data." Pp. 159-186, In *Handbook of research methods in personality and social psychology* (2nd ed.), edited by H.T. Reis and C M. Judd, Cambridge, UK: Cambridge University Press.

Weinert, Sabine, Cordula Artelt, Martin Senkbeil, Timo Ehmke, and Claus H. Carstensen. 2011. "5 Development of competencies across the life span." *Zeitschrift Für Erziehungswissenschaft*, 14(2):67–86.

Yi, Junjian, James J. Heckman, Junsen Zhang, and Gabriella Conti. 2015. "Early Health Shocks, Intra-household Resource Allocation and Child Outcomes." *The Economic Journal*, 125(588):347–371.

Chapter IV

Does Hard Work Beat Talent? The (Unequal) Interplay between Cognitive and Non-Cognitive Skills

Carlos J. Gil-Hernández

Abstract

It has long been argued that non-cognitive traits such as perseverance and motivation might outplay cognitive ability in explaining status-attainment. Cognitive and non-cognitive skills are key predictors of educational success and indicators of merit for liberal theories of equal opportunity. Nevertheless, even when accounting for socioeconomic-status (SES) inequalities in skill formation, disadvantaged pupils are less likely to make it to college. According to *compensatory theories*, SES-inequalities in educational transitions are disproportionally found among low-performing students due to status maintenance drives. However, little is known about the mechanisms accounting for this pattern. As cognitive and non-cognitive skills may be *complements* or *substitutes* in predicting educational outcomes, I test whether high-SES students compensate for low cognitive skills by high non-cognitive skills in the transition to academic upper secondary. I further contribute to the literature by exploring mechanisms such as teachers' bias and parental aspirations. I draw from the *National Educational Panel Study* to study a cohort of German students from grades 1 to 5, when early tracking is enforced. To minimise selective attrition bias and confounding, I apply inverse probability weights and school fixed-effects. I report four findings: (1) high-SES students at the same level of skills as low-SES classmates are more likely to opt for the academic track; (2) this inequality is largest among low-skilled students; (3) high-SES students are better able to compensate for low cognitive skills by high non-cognitive skills; (4) teachers' bias in grading and track recommendations, along with (over)ambitious aspirations of high-SES families, partially account for results. These findings challenge the (liberal) conception of merit as the sum of ability plus effort in assessing equality of opportunity in education.

1. Introduction

Back in the 1970s, Bowles and Gintis (1976) already argued that personality or non-cognitive traits such as perseverance and motivation might be at least as important as cognitive ability in explaining status-attainment and the persistence of inequality across generations (Heckman and Kautz 2012:457; Duckworth et al. 2012; Almund et al. 2011; Shanahan et al. 2014). Indeed, recent research shows that cognitive and non-cognitive skills are among the strongest predictors of academic performance and attainment (Smithers et al. 2018; Duncan and Magnuson 2011).

Cognitive and non-cognitive skills are rewarded by school systems and teachers when evaluating students. These skills are also key indicators of merit (e.g., talent plus effort) for liberal theories of equal opportunity (Fishkin 2014). However, even when accounting for early socioeconomic inequalities in nurturing and developing these skills (Hsin and Xie 2016; Farkas 2003), pupils from disadvantaged backgrounds are still less likely to make progress in the educational system in comparison to equally-skilled but wealthier pupils (Jackson 2013). As education is the main channel for intergenerational social mobility, this fact has important implications for status-attainment.

Post-industrial societies combine a zero-sum game of stagnant upward mobility and lack of downward social mobility of the upper-classes (Breen and Müller 2020). This phenomenon parallels the termed *wastage of talent* among disadvantaged students who do not make it to college, and the *compensatory advantage* of not gifted but privileged children that manage to get ahead in the educational system and labour market (Bukodi, Bourne and Betthäuser 2017).

Particularly, *the compensatory advantage hypothesis* posits that inequalities by socioeconomic status (SES) in accessing educational pathways leading to college are disproportionally found among low-skilled students (Bernardi and Triventi 2018). The rationale behind this hypothesis is that affluent families are particularly motivated to mobilise their extensive resources to prevent their kids from falling down the educational ladder due to risk aversion mechanisms (Goldthorpe 2007). This is particularly the case in the negative event of low scholastic ability, when the risk of downward social mobility peaks.

According to human capital theories (Heckman 2007), skills have the property of cross-productivity, so that cognitive and non-cognitive skills may be complements or substitutes in predicting learning and educational outcomes, i.e., students with high cognitive skills may complement or reinforce their already high learning capacity in combination with high perseverance. On the contrary, pupils with low cognitive skills may substitute or compensate for it by being highly conscientious—*hard work beats talent*. Similarly, kids with high cognitive skills may display low effort and progress in the curriculum anyway—*lazy but smart*.

Up to now, there is only limited research on the interplay between cognitive and non-cognitive skills in explaining educational outcomes (Light and Nencka 2019; Esping-Andersen and Cimentada 2018). Furthermore, although the literature highlights some potential mechanisms explaining why affluent students tend to avoid downward (educational) mobility (Barone et al. 2018), such as parental aspirations, perceived chances of success and economic resources, they remain largely under tested. I argue that high-SES students may be particularly able to compensate for low cognitive skills by high non-cognitive skills—*not gifted but hardworking*—so signalling to teachers their determination to get ahead in the educational system and avoid downward mobility.

A downside of most previous research on these topics is that academic skills or performance are generally measured through grade point average (GPA) as assigned by teachers (Jackson 2013), instead of using more *objective* measures of academic ability such as externally-assessed competencies or test scores (Weinert et al. 2011). Using GPA as the main indicator of academic skills rules out the possibility of teachers' bias in assessments due to students' SES—perceptions of students' ability, behaviour, and potential—as an additional mechanism accounting for the *compensatory hypothesis* (Jæger and Møllegaard 2017; Jaeger 2011).

This chapter builds upon interdisciplinary literature on social stratification, personality psychology and skill formation to provide a three-fold contribution to the state of the art: (1) examining the interplay between cognitive and non-cognitive skills in the context of the German system of early tracking; (2) testing whether low cognitive skills are substituted by high non-cognitive skills among high-SES pupils in the transition to upper secondary; and (3) exploring plausible mechanisms accounting for these patterns: teachers' bias in grading standards and track recommendations, and parental aspirations/expectations.

I draw data from the German *National Educational Panel Study* (NEPS), studying a cohort of students from the first grade of primary until the transition into secondary education in grade 5. To minimise selective attrition bias and unobserved confounding, I apply inverse probability weights and control for school fixed-effects (FE). Non-cognitive skills are measured by teachers' ratings of students' conscientiousness at grade 3. Conscientiousness is the *Big Five* personality trait most strongly related to educational performance (Poropat 2009). I measure cognitive skills with standardised competence tests in mathematics and language carried out by external evaluators at grade 4. To account for measurement error in measuring skills and approximate true ability, I carry out robustness checks in which I use several alternative measures of competencies and non-cognitive skills taken at different grades by applying a latent factor approach.

The German system of early tracking in which some states enforce binding recommendations is especially suitable for testing the compensatory hypothesis and evaluating normative theories on skills and merit (Blossfeld et al. 2016a). As teachers are supposed to recommend secondary schools on the basis of objective criteria such as academic performance and behaviour, high-SES parents may have less room to compensate if their kids are low performers at the first important crossroad for avoiding downward social mobility. Thus, the German case is a stringent test for the *compensatory advantage hypothesis* compared to educational systems without early tracking (Bernardi and Triventi 2018).

2. The German Context

The German educational system is decentralised by federal states (*länders*) (Skopek and Dronkers 2015) but all states track students at the transition between the last grade of joint primary education and secondary education, at age 10 (grade 4) or 12 (grade 6 at orientation-level schools). Primary education is characterised by a homogeneous curriculum based on mathematics and German, without ability grouping or tracking, and generally supervised by the same teacher (Ashwill 1999). Only from the third grade onwards, do teachers start to formally grade students with report cards based on exams on subjects, behaviour and classroom participation.

Generally, after primary grade 4, most pupils have access to three track-specific types of secondary schools: lower secondary (*hauptschule*), middle secondary (*realschule*), or upper secondary (*gymnasium*). *Hauptschule* and *realschule* lead to vocational training education, whereas *gymnasiums* offer the most academically-oriented education. Some states have other types of schools; for example, comprehensive schools (*gesamtschule*) were created in the 1960s by the Social Democratic Party to integrate the three-track system into one school, but nowadays they are not available in all Federal States. These schools have a uniform curriculum until grade 7, when ability grouping is applied. From grade 9, students can receive vocational training certificates or enrol in the *gymnasium* (from grade 10) depending on their previous ability group. However, in practice, comprehensive schools are considered lower rank and have not replaced the three-tier system.

During the last year of primary education, principal teachers (i.e., teaching at least one of the core subjects) recommend to families the type of secondary schools students may attend (e.g., at the end of the first semester around February or at the end of the academic year around June). According to state school laws and specific decrees,[55] teachers[56] should grant recommendations on the basis of a student's learning potential, psychological development, academic performance, and work ethic, which are mainly proxied by GPA in mathematics, German and general studies (e.g., introductory science).[57] In practice, good study, work habits and behaviour[58], perceived likelihood of success, and potential parental support are also crucial factors in the decision-making process of teachers.

Track recommendations usually take the form of a formal letter from the school principal (previously issued at the teachers' committee), and it may be discussed at a meeting between the teacher, parents and student. If the recommendation conflicts with the parental preferences, the final decision lies either with the parents, which is the case in most Länders, the secondary school, or the school supervisory authority, depending on the Federal State law (Jürges and Schneider 2007). Nonetheless, if parents finally choose a non-recommended type of school in which a positive recommendation is a prerequisite for admission, "the children usually have to pass an entrance examination and/or successfully complete a trial period in the selected school (*The Press and Information Office of the Federal Government*)." Still, compensatory strategies for school choice might be followed by high-SES parents, especially when it comes to the proactive search of alternative schools offering the *Gymnasium*, affording transportation costs if these schools are far away, and providing private tutoring to pass the entrance examination. These pragmatic issues in the process of a school search might be

[55] *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe) [Transition from primary to lower secondary schools and promotion, observation and orientation in grades 5 and 6 (so-called orientation level)].* Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany: Berlin. *https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf*

[56] It should be noted that teachers do not have formal training on how to proceed in the recommendation process.

[57] The last grade of primary has the largest weight, however, teachers take into account the whole progression of students across primary education in assessing the final grade. In some federal states, there are official GPA thresholds to grant a positive recommendation for the Gymnasium (Conference of Ministers of Education (KMK)). These thresholds range from 2.5 to 2.0. in a 1-to-6 scale where 1=very good and 6=very poor.

[58] Many states also apply behaviour-based grades (*Kopfnoten*) (e.g., orderliness).

related to the way in which parental aspirations and expectations influence track choice in Germany.

Due to these institutional arrangements, parents of children with low academic performance may have less room to influence track decisions. Thus, the German case is a stringent test for the *compensatory advantage hypothesis* compared to educational systems without early tracking. In tracked educational systems, it seems that ability-differentials by family SES account for the largest share of inequality in transition rates (Blossfeld et al. 2016a). However, there is still ample room for high-SES families to choose and deploy compensatory strategies to prevent their low-ability kids from downward mobility.

3. Theoretical Background and Previous Findings

3.1. Non-Cognitive Skills and Educational Outcomes

Since the early 2000s, research on non-cognitive skills has skyrocketed with a 400% increase in publications across social science disciplines (Smithers et al. 2018). This is mainly due to the established causal effect of non-cognitive skills on positive life outcomes, such as educational achievement, SES and health, and its potential for early interventions to reduce social inequalities (Heckman 2007). Among the most investigated non-cognitive skills, we can find the following: academic motivation, socio-emotional skills or behavioural problems, personality traits, self-regulation, locus of control, and self-esteem.

Despite this growing endeavour, non-cognitive skills are still poorly defined and measured, raising serious concerns regarding measurement overlapping, validity and error (Conti and Heckman 2014). By definition, non-cognitive skills are conceptually different from cognitive skills but, as we will see below, this is not necessarily the case.

Non-cognitive traits or skills are expressed as thoughts, feelings, and patterns of behaviour (Heckman and Kautz 2012). Personality psychologists chiefly measure these traits through rating scales, reported by the self or others,[59] that capture behaviour in real-life situations, or performance on laboratory-based tasks such as the *Marshmallow test* (Watts et al. 2018). However, these measures are not error-free and can only proxy for underlying or true psychological traits. The manifestation of traits is situation-specific (e.g., home, classroom,

[59] "Correlations between self-reports of personality traits and observers' reports of personality traits range from $r = 0.29$ to $r = 0.41$ (Killonen and Kell 2018:8)."

friends), depends on incentives, and evolves over the life course. Nonetheless, traits are stable enough across situations to claim their existence (Borghans et al. 2008). Genetic variation moderately shapes individual differences and stability in behaviour or personality, but they are not fixed at birth (Knopik et al. 2017:255; Kraphol et al. 2014).[60] The term "character skills" further accounts for sub traits or facets of personality that are thought to have more potential for malleability and early intervention, since personality traits are generally more stable and genetically influenced.

Personality psychologists have developed a well-accepted taxonomy of personality traits termed as the *Big Five domains* or *Five-Factor Model* (McCrae and Mõttus 2019; Costa and McCrae 1992a).[61] In a meta-analysis of the Big Five personality traits related to academic performance across primary, secondary and post-secondary, Poropat (2009) found that Conscientiousness is by far the largest predictor[62] of grade point average (GPA), with partial correlations over 0.2 after controlling for cognitive ability. Interestingly, some studies argue that this association between conscientiousness and GPA is as large as the one found between cognitive ability and GPA or educational attainment (Borghans et al. 2016; Duckworth et al. 2012; Heckman and Kautz 2012:457; Almund et al. 2011).[63]

Conscientiousness can be defined as "the tendency to be organized, responsible, and hardworking (American Psychology Association Dictionary), as well as "the propensity to follow socially prescribed norms for impulse control, to be goal directed, to plan, and to be able to delay gratification and to follow norms and rules (Roberts et al. 2009:369)." Conscientiousness shows high intra-individual stability—mainly driven by genetics—and it remains pretty constant at the mean-level from late preschool or mid-childhood into early adolescence, when it then steadily increases from early-adulthood due to the acquisition of more demanding adult roles (Akker et al. 2014; Eisenberg et al. 2008). Still, there is scarce longitudinal evidence on conscientiousness' development in childhood and adolescence, and the study of its change over time presents methodological challenges.

[60] "The study of the genetic heritability of sub traits has received limited attention in comparison to the Big Five. Environmental factors are also important, but their role is almost entirely driven by nonshared environmental factors. This does not necessarily mean that the family is not important, but that the relevant experiences are child-specific within the families. (Knopik et al., 2017:256)."

[61] Openness to Experience (i.e., culture), Conscientiousness, Extraversion, Agreeableness (i.e., likability, friendliness), and Neuroticism (i.e., emotional stability: locus of control and self-stem), forming the OCEAN acronym.

[62] According to Duncan and Magnuson (2011), the main socio-emotional skills related to educational outcomes are emotional regulation (ability to control anger, sadness, joy); behavioural problems (internalising: depression, withdrawn behaviour; externalising: antisocial, conduct disorders, aggression); (cognitive) self-regulation (executive function, planning, sustaining attention, task persistence, inhibition of impulsive responses); and motivation (aspirations, expectations, goals).

[63] Further evidence shows that this association is mediated by positive study habits and attitudes, and effort, among others (Almund et al. 2011:143-144).

There is lack of consensus about the hierarchical structure of the facets or sub traits comprising the Conscientiousness factor in adulthood, and about the identification of its basic architecture in early-childhood temperament (Conti and Heckman 2014). However, the most prominent sub trait or facet comprised within the Conscientiousness factor, which can be considered as having its basis in childhood, is self-regulation (also known as effortful control) in the domains of affect, activity, and attention (Eisenberg et al. 2014; Rothbart and Bates 2006:100).[64,65,66] According to Eisenberg et al. (2014:5), "Early self-control is virtually synonymous with elements of conscientiousness such as persistence, being organized, and self-discipline, and that one's ability to control attention facilitates internalization of societal values."

Self-regulation skills are key for learning due to their role in facilitating engagement and participation in academic tasks such as being able to sit, concentrating at tasks, persisting at tasks despite frustrations, and preventing disruptive behaviour in the classroom (Diamond et al. 2007). Several studies found consistent associations between these skills and academic performance at preschool and early primary education (Duncan and Magnuson 2011; Polderman et al. 2010). Other longitudinal studies controlling for baseline intelligence and grades, thereby accounting for potential omitted variables and teachers' bias, found that self-regulation remains predictive of school grades (Almlund et al. 2011:143-144; Duckworth and Seligman 2005).

3.2. The Interplay Between Cognitive and Non-Cognitive Skills

Personality traits or non-cognitive skills are, by definition, conceptually different from cognitive skills. Accordingly, most measures of personality are only weakly correlated with intelligence quotient (IQ) ($r \leq 0.30$) (Borghans et al. 2008). Cognitive skills are generally measured with IQ tests, scores on achievement tests, and GPA, although they are far from being perfectly correlated (Borghans et al. 2016). Fluid or non-verbal IQ tests were intended to capture a person's capacity to abstract reasoning and solve novel problems, independent of

[64]Effortful control includes the following domains: (a) attention-focusing (the tendency to maintain an attentional focus upon task-related channels), (b) attention-shifting (the capacity to intentionally shift attentional focus to desired channels, thereby avoiding unintentional focusing on particular channels), (c) inhibitory control (i.e., the capacity to suppress positively toned impulses and resist the execution of inappropriate approach tendencies), and (d) activation control (i.e., the capacity to perform an action when there is a strong tendency to avoid; (Eisenberg et al. 2014)."

[65] Duncan and Magnuson (2011:12): "Self-regulation has been defined as the processes by which the human psyche exercises control over its functions, states, and inner processes. It involves the ability to evaluate the steps and actions required to meet a desired goal and to control behavior deliberately in order to reach that goal."

[66] The four sub traits or facets comprised within the Conscientiousness factor most reliably identified in previous work are: responsibility (e.g., punctuality), self-control (e.g., ignoring distractions), industriousness (e.g., perseverance, pursuing goals to their ends) and orderliness (e.g., keeping an organised work space) (Eisenberg et al. 2014).

previous numerical and verbal knowledge (Nisbet et al. 2012). In turn, achievement tests were designed to blindly evaluate domain-specific cognitive competencies that are crucial for success in school and later life, such as reading, mathematical or scientific literacy, in a more objective way than grades assessed by teachers (Weinert et al. 2011).

Indeed, conscientiousness and IQ correlations are small in magnitude (Murray et al. 2014). However, response to cognitive tests also reflects context-specific non-cognitive skills such as motivation to perform, persistence and effort. Hence, low-stakes standardised tests of academic competencies, such as the *Programme for International Student Assessment* (PISA), that were thought to be independent of non-cognitive skills–and teachers' bias in evaluating them-also capture them to a certain extent (Azzolini et al. 2019; Kyllonen and Kell 2018:9). This is especially the case for the GPA, displaying the highest correlation with personality traits[67] (Borghans et al. 2016; Duckworth et al. 2012). Similarly, self-reported personality assessments may also reflect cognitive factors in the understanding of personality descriptions.

Beyond problems of measurement cross-contamination, cognitive and non-cognitive skills do not work independently,[68] and their association seems weak-to-moderate, but generally positive. Human capital theories posit that cognitive and non-cognitive skills have the property of *cross-productivity* (Heckman 2007), referring to the virtuous circle or positive feedback loops in their development. For instance, there is a well-documented positive correlation between self-control skills and IQ (Rikoon et al. 2016). It has been found that children with extremely high IQs also score very high on self-control (Calero et al. 2007). Likewise, Meldrum et al. (2017) found a positive longitudinal association between intelligence and self-control, accounting for prior self-control, child executive functioning, maternal intelligence, and maternal self-control.

Alternatively, the *intelligence compensation hypothesis* (ICH) posits a negative association between cognitive and non-cognitive skills (Rammstedt et al. 2018), or the idea that "Conscientiousness acts as a coping strategy for relatively less intelligent people" when striving for similar achievement as higher-ability peers (Rikoon et al. 2016:21; Murray et al.

[67] The correlation between conscientiousness and standardised tests is lower than the one with school grades: "IQ predicted changes in standardized achievement test scores over time better than did self-control, whereas self-control predicted changes in report card grades over time better than did IQ (Duckworth et al. 2012)." Duckworth et al. (2012) concluded that "intelligence may influence an individual's ability to learn and solve problems independent of whether or not they receive instruction, yet self-control facilitates achievement by contributing to an individual's ability to study (a task which requires focus and allocation of time), complete tasks and assignments, and the tendency to take an active role in classroom participation."

[68] Duncan and Magnuson (2011) point out that the widespread term of non-cognitive skills is misleading given that it contains a wide array of traits, such as attention control, which is an inherently cognitive task regulated by the executive functions of the brain (e.g., emotional, attentional, and inhibitory control).

2014). By the same token, the ICH argues that "individuals higher in cognitive ability are proposed not to increase in conscientiousness because their higher cognitive ability allows them to accomplish more with the same or less effort (Murray et al. 2014:17)." However, as shown by Murray et al. (2014), this proposed negative IQ-conscientiousness association is explained by positive sample selection (e.g., samples of college graduates; collider bias), or high-stakes testing settings in which maximum effort is exerted.

Given previous findings, we can conclude that cognitive and non-cognitive skills are positively associated. However, the causal direction of their positive association is far from being clear (Metcalfe et al. 2013; Bub et al. 2007). According to Nisbett et al. (2012:151), there are at least three plausible explanations: "(a) The ability to self-regulate could be a manifestation of intelligence; (b) these constructs could share common variance such that they are both affected by a third variable[69] [e.g., genes; executive functions, parental SES] (Malanchini et al. 2020, 2018; Uka et al. 2019; Fleming et al. 2016); (c) self-regulation could be one of the processes that facilitate the development of intelligence."

Independently of the direction of the association between cognitive and non-cognitive skills and their common or causal foundations, it is clear that they have unique sources of variation and represent different latent skills (Malanchini et al. 2020, 2018). The bottom line is that it is imperative to adjust for non-cognitive skills when assessing the effect of cognitive skills, and vice versa—ideally using a latent factor approach to account for measurement error and constructs overlapping.

Apart from being correlated, skills may also interact by being complements or substitutes in the production of (self-regulated) learning and educational outcomes (Light and Nenka 2019): students may affect their own learning by selecting how to mix their skills as a function of their current information or metacognition. For instance, students with high cognitive skills may complement or reinforce their already high learning capacity by combining with high perseverance—*self-regulated learner*—thus yielding larger returns on perseverance in test scores or GPA than low-ability kids.

On the contrary, kids with high cognitive skills may display low effort and achieve educational milestones anyway—*lazy but smart.* Similarly, pupils with low cognitive skills may find it necessary to substitute or compensate for it by being highly conscientious in order to

[69] Twin studies showed that, in childhood, variation in general cognitive ability is genetically correlated with variation in attention problems, pointing to a shared set of genes that influence both IQ at age 12 and attention problems at age 5 (Polderman et al. 2006:204).

progress in the educational system—*hard work beats talent*. Whether and how the effect of non-cognitive skills on educational outcomes differs across the distribution of cognitive skills is an open empirical question (Light and Nenka 2019:142-143).

To my knowledge, Light and Nenka (2019) is the only available evidence assessing the interplay of cognitive and non-cognitive skills on educational attainment (high school graduation; college enrolment). They found that the largest marginal effect of non-cognitive skills, as measured by grit, is concentrated at the bottom and top extremes of the cognitive ability distribution. Thus, these authors argue that high-cognitive-ability students adopt self-regulated learning where their grit is especially productive in challenging tasks, while for low-cognitive-ability students, grit seems to play a compensatory role in learning. Whether these patterns of skill complementarities and substitution vary by parental SES remains unknown.

3.3. The Interplay Between Skills and Parental SES

It is well-known that the level of development of cognitive and non-cognitive (Attanasio et al. 2020) skills varies by parental SES already from preschool age—known as primary effects (Jackson 2013)—due to parenting, genetics, health, and pre-school quality, among other factors (Passaretta et al. 2020; Hsin and Xie 2016; Duncan and Magnuson 2011). Likewise, it is well-established that high-SES families have a systematic advantage in transition rates to higher educational levels in comparison to low-SES families, even when controlling for skill differentials (secondary effects; Jackson 2013). However, much less is known on whether (and how) parental SES and skills may interact in predicting early educational outcomes.

Socioeconomic resources and skills may be complements or substitutes when it comes to predicting status attainment (Damian et al. 2015). On the one hand, The *Matthew effect hypothesis* predicts that the "rich get richer" so that skills are the strongest predictors of attainment among high-SES families. On the other hand, in line with the *resource substitution hypothesis*, low-SES students might overcome their background disadvantage by relying on high personality or cognitive skills, while skills may be less predictive of status-attainment for high-SES students, who can compensate with high resources (Liu 2019).

In a similar vein as the *resource substitution hypothesis*, *the compensatory advantage hypothesis* also predicts that academic skills are less predictive of educational outcomes among high-SES students (Bernardi 2014; Bernardi and Cebolla 2014). Drawing from rational action theories (Breen and Goldthorpe 1997), the logic behind this idea is that affluent families are particularly motivated to mobilise their extensive resources to prevent their kids from falling down the educational ladder due to risk aversion to social demotion (Goldthorpe 2007). This is particularly the case in the negative event of low scholastic ability, when the risk of downward social mobility is at a maximum (Bernardi and Grätz 2015).

The implication of this hypothesis is that inequalities by SES in educational transitions are disproportionally found among low-skilled students. By contrast, low-SES families with similar abilities to more advantaged children would systematically follow less *ambitious* educational paths due to their lower level of resources, risk aversion to downward mobility, and expected chances of success (Barone et al. 2018). From a normative standpoint, the notion of *compensatory advantage* is at odds with ability being the main criterion of merit for evaluating equality of opportunity in education (Fishkin 2014).

Indeed, previous research found that for those low-SES children who lack cultural, economic and social resources, both cognitive (Holm et al. 2019: GPA; Bernardi and Triventi 2018: GPA; Bernardi and Cebolla 2014: GPA; Damian et al. 2015: test scores in mathematics, and verbal and spatial ability) and non-cognitive skills are more predictive of education and status-attainment than for high-SES children (Liu 2019: index on approaches to learning, self-control and interpersonal skills; Esping-Andersen and Cimentada 2018: ambition, perseverance and discipline; Damian et al. 2015: Big Five personality traits; Shanahan et al. 2014: Big Five).

Shanahan et al. (2014) reported that low-SES children compensate for their lack of resources by being conscientious, agreeable, emotionally stable, and open in the attainment of education, wages and occupation. Damian and colleagues (2015) replicated the former study finding that, after controlling for cognitive skills as measured by scores in mathematics, and verbal and spatial abilities, the interaction between parental SES and non-cognitive skills became insignificant, but the interaction between SES and cognitive skills remained solid. Accordingly, Damian et al. (2015) argued that cognitive skills have a greater compensatory effect *vis-a-vis* non-cognitive skills among low-SES children.

Unfortunately, research on whether and how cognitive skills, non-cognitive skills and parental SES interact to affect students' early educational outcomes is scarce. Some studies analysed the interaction between (1) parental SES and cognitive skills (Holm et al. 2019; Damian et al. 2015; Bernardi 2014); (2) parental SES and non-cognitive skills (Liu 2019;

Damian et al. 2015; Shanahan et al. 2014); and (3) cognitive and non-cognitive skills (Light and Nenka 2019) in predicting status-attainment.

However, to my knowledge, up until now only Esping-Andersen and Cimentada (2018) have looked at the joint interplay between all three of them. They concluded that (1) the shelter against downward mobility enjoyed by the children of advantaged origins cannot be attributed to skills' compensatory effects; and (2) that the chances of upward mobility for disadvantaged children are maximised under the combination of high cognitive and non-cognitive skills (*skills complementarities*). However, their measures of skills drawing from PIAAC data raise some concerns on validity and reverse causality, as they were captured in adulthood after education and SES were achieved. Finally, but no less importantly, among these reviewed studies, only Liu (2019) looked at early educational outcomes—test scores or academic achievement in adolescence—in the process of social stratification.

Consistently with the *compensatory advantage* (*resource substitution*) hypothesis, I posit that *skills are less (more) predictive of educational outcomes for high-SES (low-SES) students.* Moreover, in line with the notion of *skill substitution*, I hypothesise that *high-SES families are particularly able to compensate for low cognitive skills by high non-cognitive skills*, so signalling to teachers their determination to get ahead in the educational system.

3.4. Mechanisms

Although the previous literature highlights some potential mechanisms[70] explaining why wealthy students tend to get ahead in the educational system in comparison to equally-skilled but worse-off counterparts (secondary effects), they remain largely under applied to test the *compensatory hypothesis* (Bernardi 2014). In the remainder of this chapter, I elaborate on plausible mechanisms that may account for my predictions on *compensatory advantage* and *skill substitution* in the German context of early ability tracking—namely, teachers' bias in assessments and parental (over)ambitious aspirations/expectations.

3.4.1. Teachers' Bias

Teachers are not free of bias in their evaluation of students' academic abilities, classroom behaviour, and educational expectations (Alesina et al. 2018; Esser 2016). Teachers, as all

[70] Economic resources (direct and indirect costs), perceived chances of success, perceived benefits of education, risk aversion to downward mobility (Barone et al. 2018), ambitious aspirations and expectations (Zimmermann 2020), and teachers' bias—also known as *tertiary effects* (Jæger and Møllegaard 2017; Esser 2016).

human beings, are exposed to implicit (subtle) biases in their judgement and behaviour. Implicit bias can be defined as "unconscious attitudes, reactions, stereotypes, and categories that affect behaviour and understanding (Boysen et al. 2009)." In the school system, implicit bias stands for unconscious gender, racial or socioeconomic bias towards students. According to the *Yale Poorvu Center for Teaching and Learning*, "Instructors can hold assumptions about students' learning behaviours and their capability for academic success which are tied to students' identities and/or backgrounds, and these assumptions can impede student growth."

According to *cultural reproduction theories* (Bourdieu and Passeron 1990), this bias is the result of teachers positively evaluating those children socialised in the dominant culture of the upper-classes, to which teachers themselves belong. Thus, *cultural reproduction* authors theorise that teachers misconceive cultural capital as academic brilliance and, as a corollary, the educational system functions as an institution of reproduction of inequality (Jaeger and Mollegaard 2017). To date, it is not clear-cut, though, whether unconscious bias (e.g., out-group bias in cognition) outweighs conscious bias (e.g., explicit racism or classism; *statistical discrimination*) in shaping teachers' judgements.

Previous research supports the existence of teachers' bias in grading standards and track recommendations as a function of students' ascribed characteristics such as ethnic and socioeconomic background (Wenz and Hoenig 2020; Triventi 2019; Geven et al. 2018; Marcenaro-Gutiérrez and Vignoles 2015; Timmermans et al. 2018, 2015; Boone and Van Houtte 2013). Teachers' bias in grading is generally measured as the difference between school grades assigned by teachers and blindly-assessed standardised test scores.[71] Even when it is difficult to infer discrimination from these empirical regularities, if teachers favour characteristics of the student or his/her home environment that are related to school success, but not associated with a child's cognitive and non-cognitive skills, it can certainly be considered as a form of bias against low-SES families (Geven et al. 2018:13-15).

As I pointed out above, low-SES parents may be especially sensitive to distorting biases in the signalling information that grades provide (Martínez de La Fuente et al. 2020), likely pushing their educational expectations downwards. This distorting effect can be reinforced when low-SES students are just around the minimum GPA cut-off to grant access to the

[71] The correlation between GPA and test scores is far from perfect at about 0.63 (Südkamp et al. 2012). Thus, it is crucial to control for previous students' non-cognitive skills when assessing teachers' bias as the difference between GPA and test scores. Otherwise, this measure could be just reflecting the fact that students tend to exert less effort in low-stakes testing settings.

academic track, where information on potential success is particularly unclear (Holm et al. 2019; Batruch et al. 2018; Bernardi and Cebolla 2014).

Furthermore, given that the *Gymnasium* is considered as a very demanding school type, parents tend to be interested and involved in their children's academic progress to a greater extent than other secondary schools (Ashwill 1999). Thus, if teachers presume that low-SES parents will not be able to support their low-ability kids during the *Gymnasium*, they may be reluctant to grant a positive recommendation in anticipation of a low likelihood of success (Krolak-Schwerdt et al. 2018). Following the concept of *statistical discrimination*, teachers may issue track recommendation relying on SES-stereotypes and their past experience on SES-average chances of success in the academic track under imperfect information on potential outcomes.

On the contrary, teachers may perceive that low-ability, but striving, kids from high-SES families will have parental support and a fair likelihood of success in the demanding setting of the academic track (Krolak-Schwerdt et al. 2018). High-SES parents may be especially persuasive in influencing teachers' tracking decisions thanks to their high level of cultural capital and information on the educational system (e.g., parent-teacher meetings; participation in school council; volunteering at school) (Forster and van de Werfhorst 2020; Lareau 2015).[72] High-SES parents may be especially prone to meeting with the teachers to monitor their kids' progress during elementary school and arrange a consultation meeting with the teacher in case of a negative formal recommendation. Moreover, teachers might anticipate that high-SES parents will complain about non-academic recommendations, and therefore give academic recommendations to their low-ability kids just to avoid any awkwardness (Barg 2012).

3.4.2. Parental Aspirations and Expectations

It is well-known that high-SES parents have higher educational aspirations (*ideal*) and expectations (*realistic*) for their children than low-SES parents (Morgan 1998), even when ability is held constant (Zimmermann 2020). Generally, SES-differences in parental educational aspirations tend to be narrower than educational expectations, as the former express idealistic goals in a world without constraints. Thus, I believe that the level of downgrading from college aspirations to expectations can tell us more about how families

[72] A parent from Eastern Germany expressed the following statement with regard to the complexity of the tracking decision (Ashwill 1999:87-88): "Parents do not feel competent to make all these important decisions for their children at such an early age. I think that it is very hard to know what talents and inclinations children will develop later in life."

from different socioeconomic backgrounds interpret the cost (indirect costs and opportunity costs), benefits (short-term earnings' returns) and necessary ability to pursue academic education. Low-SES families may be more conservative and make less risky choices, whilst the high resources and relative risk aversion to downward mobility of high-SES families might make them more insensitive to the signal of low ability.

For all of the above, I conjecture that (over)ambitious expectations of high-SES parents for their low-ability kids may pressure teachers to give higher grades and positive recommendations in the course of primary education (e.g., parent-teacher meetings) (Barg 2012). Besides, high-SES parents may directly ignore bad grades or a negative recommendation by enrolling their kids in the Gymnasium in those federal states without binding constraints or prepare their kids for the entrance examination or trial period in those states with binding regulation—i.e. likely with the extra help of private tutoring (Jürges and Schneider 2007). I further contend that pupils with low cognitive skills coming from high-SES families, at least do need to exert high effort or display a certain minimum level of non-cognitive skills to signal to teachers (and their parents) that they will have a fair chance of success in the demanding setting of the academic track.

4. Data, Variables and Methods

4.1. Data

I draw data from Starting Cohort 2 (SC2) of the NEPS(Blossfeld et al. 2011).[73] This study initially surveyed 4-year-olds reaching school age in 2012 in a representative sample of day-care facilities in the first wave (2010/2011; n=2,949), to be followed up in the second year of kindergarten (wave 2) and into the school context (n=557 in wave 3). Starting Cohort 2 also comprises an augmentation or refreshment sample of students entering grade 1 of primary education in school year 2012/2013 (n=6,176) in wave 3, sharing the elementary school context with the former Kindergarten group.

[73] This inquiry uses data from the NEPS: Starting Cohort Kindergarten, 10.5157/NEPS:SC2:8.0.0. From 2008 to 2013, NEPS data were collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

Table 1. Variables by wave, grade and age

Variables	**Wave 3 Grade1**	**Wave 4 Grade 2**	**Wave 5 Grade 3**	**Wave 6 Grade 4**	**Wave 7 Grade 5**
	2012/2013	2013/2014	2014/2015	2015/2016	2016/2017
	Age 6-7	Age 7-8	Age 8-9	Age 9-10	Age 11-12
Socio-demographics					
Parental SES	X				
Migration background	X				
Gender	X				
School ID	X				
Non-cognitive skills					
Conscientiousness			**X**		
Concentration / persistence ability	**X**	**X**	**X**	**X**	
Readiness for exertion	**X**				
Joy of learning	**X**				
Behavioural problems: pro-social behaviour (SDQ)			**X**		
Behavioural problems: peer Problems (SDQ)			**X**		
Behavioural problems: disruptive behaviour (TASB)		**X**			
Cognitive skills					
Cognitive basic skills		X			
Domain-specific competencies:					
Maths	X	X		X	
Scientific literacy	X		X		
Vocabulary	X		X		
Orthography				X	
Reading literacy		X		X	
Reading speed		X			
Outcomes					
Type of school attended					X
Parental aspirations / expectations	X			X	
Teachers' track recommendation				**X**	
Last year's annual GPA in maths and German					**X**

Notes: in bold=assessed by teachers; Joy of learning (1-4; 2 items): (i) *Child enjoys going to school*, (ii) *Child enjoys learning at school*; Readiness for exertion (1-4; 4 items): (i) *Child works carefully with the work materials*, (ii) *Child makes an effort when assignments are difficult*, (iii) *Child gives up easily if something is difficult*, (iv) *Child works diligently in class.* Concentration/persistence (1-5; 1 item): (i) *Persistence and ability to concentrate (e.g., remaining occupied with something for a longer period of time) [compared with other children of the same age]* ; Strength and Difficulties Questionnaire (SDQ): peer relationship problems (total sum (0-10) of 5 items (0-2)): (i) *Loner: Mostly plays alone*, (ii) *Has at least one good friend*, (iii) *Generally popular with other children*, (iv) *Is teased or victimised by others*, (v) *Gets along better with adults than with other children*; SDQ: pro-social behaviour (total sum (0-10) of 5 items (0-2)): (i) *Considerate*, (ii) *Likes to share with other children e.g., sweets, toys, crayons etc.*, (iii) *Likes to help when others are hurt, ill or upset*, (iv) *Kind to younger children*, (v) *Often helps others voluntarily, e.g., parents, teachers or other children*; Teacher Assessment of Social Behaviour (TASB) (total sum (3-15) of 3 items (1-5)): (i) *Disturbs other children in their activities*, (ii) *Takes a predominant position within the group*, (iii) *Interrupts other children.*

The refreshment sample was followed up from elementary education to the transition into secondary education as of wave 7 (2016/2017) onwards. Analyses use from waves 3 (participation rate of 97.4%) to 7 (participation rate of 56%) of the augmentation sample. This sample was drawn based on a nationwide representative sample of students at elementary schools following a two-stage approach –schools were drawn in the first step, and students in the second step. Students were interviewed and tested in the classroom context via paper-and-pencil questionnaires (PAPI) after parental consent. Parents who accepted to participate, in addition to giving permission to their kids, were questioned through computer assisted-telephone interviews (CATI). Class teachers (PAPI), and school principals (PAPI) were also invited to participate in the study to report information on themselves, students and classroom/school characteristics (Steinhauer et al. 2014:42).

4.2. Variables

Table 1 displays the main variables used in this study by survey wave, grade and students' age.

Parental SES. Parental SES is proxied with the highest parental *International Socio Economic Index of Occupational Status* (ISEI-88) measured in wave 3 (assuming it is time-constant)[74] and codified into a dummy capturing low and high SES (0=q1-q2, and 1=q3-q4) (Ganzeboom and Treiman 1996). I operationalised parental SES into a dummy to maximise sample size to carry out moderation analysis by SES subgroups. I carried out a robustness check with alternative measures of parental SES such as household income (Low: 0=q1-q2; High: 1=q3-q4) and highest parental education (ISCED-97; at least one parent with a college education or above vs rest), and the results hold (see Appendix Tables A.2.–A.3. for a replication of the main models on track choice).

Socio-Demographic Controls. All models control for time-constant socio-demographics measured in wave 3, namely, gender and migration background (1=native origin; 2=first generation migrant; 3=parents born abroad and children born in Germany (generation 2.0)) to estimate the effect of parental SES, net of migration status.

Basic Cognitive Skills. Fluid intelligence or non-verbal cognitive ability is proxied in wave 4 at grade 2 through the *Picture Symbol Test* (NEPSBZT) assessing perceptual speed (two sets

[74] A reasonable assumption given its high intra-class correlation of 0.73 and low within-individual standard deviation of 6.3.

of 21 items on matching figures or numbers with graphical symbols) and the matrices test (NEPS-MAT) assessing figural reasoning (one set of 12 items completing fields in matrices of geometrical symbols). As recommended by NEPS technical reports (Haberkorn and Pohl 2013), I estimated a standardised average between both tests to proxy for non-verbal IQ and reliably capture domain-specific competencies later on. Since I can only observe basic cognitive skills at wave 4, I assume basic cognitive skills to be time-constant. Previous research argues that non-verbal IQ displays high intra-personal correlation in the short-term and that its main sources of stability over time are genetic in origin (Briley and Tucker-Drob 2013; Deary and Johnson 2010).

Non-Cognitive Skills. Non-cognitive skills are mainly measured by teachers' ratings of the students' conscientiousness (one of the Big Five personality traits) in wave 5 at grade 3.[75] Following the "Fünf Faktoren Fragebogen für Kinder—Kurzform" (FFFK-K, Five Factor Questionnaire for Children—Short Form) instrument (Blossfeld et al. 2016b), a short and age-adapted version of the "Big Five bipolar adjective scales" (Asendorpf and van Aken 2003), two facets on conscientiousness are asked to teachers: disorganised/organised (0-10); and focused/distractible (0-10). This measure has acceptable reliability at $\alpha=0.7$ and validity (Weinert et al. 2007). Additional analyses drawing from the longitudinal subsample of Kindergarten kids shows that conscientiousness is relatively stable over time as measured at grade 2 of Kindergarten (wave 2) and grades 2-3 of primary (waves 4-5), with correlations of around 0.5. It should be noted, though, that teachers' evaluations of students' behaviour may be subjected to bias by students' SES, so that high-SES students may be over-evaluated. If this were so, the effect of parental SES on track outcomes, net of skills, would be underestimated.

Domain-Specific Cognitive Skills. I measure domain-specific cognitive skills with low-stakes competence tests (test scores) on mathematics (24 items) and language (reading literacy with 33 items; and orthography with 37 items in two tests) administered by external evaluators and supervised by school teachers at grade 4 in wave 6.[76] These tests follow a similar methodology as large-scale international assessment studies (e.g., PISA) (Weinert et al. 2011). Test scores are designed following *Item Response Theory* and provided by NEPS as

[75] Robustness checks using conscientiousness as reported by parents in wave 4 at grade 2, or controlling for it as a lagged variable in addition to the teachers' report, yields consistent results.

[76] In line with the theoretical discussion and previous evidence summarised above in *sections 3.1.-3.2.*, I assume that personality traits and non-verbal IQ are antecedent skills to performance in blindly-evaluated and low-stakes test scores. I also assume that performance in test scores is not causally associated with latter changes in baseline personality traits or non-verbal IQ.

weighted maximum likelihood estimates (WLEs) or sum scores. To construct a composite competence measure in grade 4 comprising mathematics and language skills, (1) I standardised these domain-specific measures within wave 6 to express students' relative position in the age-specific distribution of each domain; and (2) applied factor analysis to estimate the weighted mean z-scores across language and mathematics domains according to the first factor (only one factor retained with Eigenvalue at 2.86 and α=0.87).

Skills' Composites. The main measures of cognitive (wave 6; mathematics and language competencies) and non-cognitive skills (wave 5; conscientiousness) studied in this chapter are captured at a single time point with single measures. This strategy induces measurement error bias such as attenuation, since teachers tend to evaluate the whole progression of students across elementary education, not just snapshots, and performance in low-stakes standardised tests and teachers' reports of students' behaviour may not capture true ability.[77] To account for measurement error and approximate true ability, I constructed skills' composites by relying on several measures of cognitive competencies (13 tests on language, mathematics and scientific literacy) and non-cognitive skills (25 items on motivation, effort, behavioural problems, attention/persistence skills, and personality) taken across grades 1-to-4 (see table 1 below for details on the variables). In Appendix A.1. on measurement error, I explain in depth the technical details of the process followed to construct the skills' composite measures. I replicated the main analyses with these alternative skills' composites as a robustness check that I comment on throughout the discussion section when relevant.

Track Choice. Track choice is the main dependent variable of this study. It is measured through the current type of school attended in wave 7,[78] operationalised through a dummy mainly distinguishing between vocational training (VT) tracks (0=*Hauptschule* and *Realschule*), and the academic track leading to college after passing the high-stakes *Abitur* exam (1=*Gymnasium*). Regarding comprehensive schools (*Gesamtschule*; around 10%), which integrate both vocational and academic tracks within the same school across grades 5-to-10 (implementing ability grouping from grade 7), I consider them as lower-rank and include them in the denominator together with VT tracks due to their lower ability and SES composition with respect to the academic track (0=*Hauptschule*+*Realschule*+*Gesamtschule*). I carried out a

[77] For some non-cognitive measures, parental and students' reports are available. I additionally built a factor of non-cognitive skills taking parental and students' reports into account and the results hold. However, I decided to only use teachers' reports as the main measure as they better capture children's behaviour in the classroom, which is the focus of this work.

[78] I additionally update the school attended by drawing information from wave 8 among those observations that change school during these waves (most cases upgrading into secondary schools due to grade retention) or were missing in wave 7.

robustness check by excluding comprehensive schools and the results hold. If students were not held back during primary (1.5%), by wave 7 (course 2016/2017), they should be enrolled in grade 5, when tracking starts in all German Federal states except for the following Eastern Länders in which primary lasts until grade 6 (and tracking is applied from grade 7): Berlin, Brandenburg and Mecklenburg-Western Pomerania (Skopek and Dronkers 2015). Accordingly, I exclude those students still attending primary schools in wave 7 in Eastern Germany, or orientation-level schools (*Förderschule*; 0.8%) that also delay tracking until grade 7. Special needs schools and schools for gifted children are also excluded from the analysis.

Grade Point Average. In wave 7, students are asked for their last year's annual school report in German and mathematics, corresponding to school year 2015/2016 at grade 4. In elementary school, GPA is provided in a 1-to-6 scale in which 1=very good; 2=good; 3=satisfying; 4=sufficient; 5=inadequate; and 6=unsatisfactory. I averaged both German and mathematics GPA in their natural scale, with a Cronbach alpha at 0.77.

Track Recommendation. In wave 6, students attending grade 4 in those federal states tracking students from grade 5 are recommended a school type by their teachers. Recommendations take the form of a formal letter from the school and/or a consultation in a parent-teacher meeting. I operationalised track recommendations as a dummy in which 1=Gymnasium and 0=other types of schools, excluding those observations reporting that there was no recommendation yet (13%; held back; still attending 5-6 grade of primary). In the models predicting track recommendation, I control for the type of recommendation issued (1=consultation; 2=recommendation; 3=recommendation and consultation).

Parental Aspirations and Expectations. Parental educational aspirations for their children are measured through the following question: *No matter which school <target child's name> is currently attending or how good his grades are: What school-leaving qualification would you like him to obtain?* This question captures goals that parents can have for their kids in a world without constraints. To measure educational expectations (e.g., realistic aspirations), parents are asked the following question: *And considering everything you know now: What qualification will <target child's name> actually leave school with?* This question proxies for parental goals that their children can 'realistically' reach, given existing barriers and constraints (Morgan 1998). I decided to use parental aspirations/expectations due to the fact that, at ages 6-10, children have a low level of agency and cognitive maturity to understand their place in the social structure, to express realistic career goals, and to be able to disentangle between idealistic and realistic educational aspirations (Zimmermann 2020; Ashwill 1999). I operationalised aspirations and expectations into dummies in wave 6 at grade 4[79] distinguishing between *Abitur*=1 (*general university entrance qualification* after *Gymnasium*), and *other types of school-leaving qualifications*=0.

4.3. Attrition and Weighting

To minimise selective attrition bias from wave 3, I generate inverse probability weights (IPW) to adjust for attrition in the multivariable analyses (Skopek and Passaretta 2018). Attrition rates in waves 4 (8.5%), 5 (15.9%) and 6 (10.6%)[80] are generally low. However, from wave 6 onwards, students move to the individual field (wave 7), as this is the first year in which transition to secondary schools is possible, meaning that students and parents are jointly surveyed and tested at home. Thus, the attrition rate is considerable larger in wave 7 (43.7%) with respect to previous waves.

I model attrition with a set of socio-demographic and skills' predictors, namely: gender, migration background, parental SES, and cognitive and non-cognitive skills, which correspond to the analytical variables of this inquiry. Particularly, I run logistic models to predict the probability of attrition from wave 3 to 6, and from wave 6 to 7, assigning the inverse probability of attrition to observations depending on attrition status. Finally, I generate longitudinal weights by multiplying the design cross-sectional weight of wave 3, times the conditional IPW predicted from waves 3 to 6, times the conditional IPW predicted from waves 6 to 7. As a robustness check, I additionally adjusted analyses for longitudinal weights as provided by NEPS, which model the participation of children (and jointly with their parents) at each wave transition across waves 3-to-7, to the multivariable models and results are highly consistent with the IPW I generated.

4.4. Sample Selection and Missing Values

Table 2 summarises the share of missing observations due to item non-response within the sample of students that participated in all waves of the refreshment sample of SC2 (waves 3-7; n=2,806), so deleting temporary dropouts. I assumed missing at random and applied list-wise deletion to the variables of interest in the analyses. Accordingly, four analytical samples

[79] Aspirations (and expectations) are highly stable over time: 83% (75%) of low-SES families and 88% (82%) of high-SES families do not change their aspirations (expectations) between grade 1 and grade 4.
[80] This increase with respect to wave 5 is due to the inclusion of temporary dropouts not observed in wave 5.

were selected with the following inclusion criteria: (1) participation in all waves from wave 3-to-7; and (2) non-missing values on type of school attended in wave 7. From this baseline sample, I applied list-wise deletion to each of the four outcome variables studied (track decision; GPA; track recommendation; and parental aspirations/expectations) and their corresponding control variables independently. Consequently, main analytic samples (analytic samples using skills' factors) stand at 2,055 (2,680) for track choice; 1,774 for GPA (2,298); 1,458 (1,947) for track recommendation; and 1,780 (2,371) for parental aspirations/expectations.

Table 2. Missing data among participants in waves 3-to-7

	n missing	% missing
Gender (wave 3 / grade 1)	0	0
Age (waves 3-6 / grades 1-4)	0	0
Migration background (wave 3 / grade 1)	124	4.42
Parental ISEI (wave 3 / grade 1)	146	5.20
Household income (wave 3 / grade 1)	137	4.88
Parental education (wave 3 / grade 1)	128	4.56
Parental educational aspirations (wave 3 / grade 1)	297	10.58
Parental educational expectations (wave 3 / grade 1)	431	15.36
Basic cognitive skills (wave 4 / grade 2)	332	11.83
Conscientiousness (wave 5 / grade 3)	650	23.16
Competencies in maths and German (wave 6 / grade 4)	164	5.84
Annual GPA in maths and German (wave 7 / grade 4)	144	5.13
Parental educational aspirations (wave 6 / grade 4)	525	18.71
Parental educational expectations (wave 6 / grade 4)	530	18.89
Teachers' track recommendation* (wave 6 / grade 4)	525	18.71
Non-cognitive skills composite (waves 3-6 / grades 1-4)	0	0
Cognitive skills composite (waves 3-6 / grades 1-4)	1	0.04
School type/track currently attended (wave 7 / grade 5)	44	1.57
Longitudinal weight (waves 3, 6 / grades 1, 4)	264	9.41
n	2,806	

Notes: *Not included those cases with no recommendation/consultation (n=318)

Table 3. Summary statistics by Parental ISEI (analytical samples)

	Low-ISEI (q1-q2)					High-ISEI (q3-q4)				
	Mean	**SD**	**Min.**	**Max.**	**n**	**Mean**	**SD**	**Min.**	**Max.**	**n**
Age (wave 3)	7.12	0.40	5.58	8.58	926	7.04	0.37	5.42	8.33	1,129
Age (wave 4)	7.76	0.39	6.25	9.08	926	7.69	0.35	6.08	8.92	1,129
Age (wave 5)	8.77	0.39	7.08	10.08	926	8.69	0.36	7.08	9.92	1,129
Age (wave 6)	9.79	0.39	8.08	11.17	926	9.71	0.35	8.08	10.92	1,129
Female (wave 3)	0.50		0	1	926	0.49		0	1	1,129
Native background (wave 3)	0.76		1	3	926	0.85		1	3	1,129
Migration background: 1.0 Generation	0.02		1	3	926	0.01		1	3	1,129
Migration background: 2.0 Generation	0.22		1	3	926	0.13		1	3	1,129
Parental ISEI (wave 3)	43.34	9.58	1	54	926	69.73	8.82	55	90	1,129
Household income (wave 3)	3,112.38	1,415.13	500	30,000	926	4,653.79	2,714.13	1,000	100,000	1,129
Parental highest education: College (wave 3)	0.13		0	1	926	0.67		0	1	1,129
Parental educational aspirations: *Abitur* (wave 3)	0.69		0	1	620	0.86		0	1	838
Parental educational expectations: *Abitur* (wave 3)	0.54		0	1	620	0.78		0	1	838
Z - Basic cognitive skills (wave 4)	0.00	0.99	-3.70	2.67	926	0.04	0.98	-3.18	2.67	1,129
Z - Conscientiousness (wave 5)	-0.05	1.00	-2.21	1.62	926	0.18	0.95	-2.21	1.62	1,129
Z - Competencies in maths/language (wave 6)	-0.18	0.98	-3.93	3.17	926	0.30	0.95	-2.94	3.65	1,129
Parental educational aspirations: *Abitur* (wave 6)	0.67		0	1	759	0.85		0	1	1,021
Parental educational expectations: *Abitur* (wave 6)	0.59		0	1	759	0.80		0	1	1,021
Track recommendation: Academic track (wave 6)	0.57		0	1	620	0.77		0	1	838
Recommendation type: Recommendation (wave 6)	0.19		1	3	620	0.21		1	3	838
Consultation	0.18		1	3	620	0.17		1	3	838
Recommendation and consultation	0.63		1	3	620	0.61		1	3	838
Last year's Annual GPA in maths/German (wave 7)	2.27	0.80	1	6	782	1.97	0.75	1	6	992
Academic track (wave 7)	0.45		0	1	926	0.71		0	1	1,129
Z - Non-cognitive skills composite (waves 3-6)	-0.06	0.98	-2.61	2.41	1,176	0.28	0.95	-3.06	2.41	1,504
Z - Cognitive skills composite (waves 3-6)	-0.19	0.96	-2.65	3.85	1,176	0.35	0.95	-2.34	3.49	1,504
Longitudinal weight (waves 3,6)	3.28	2.50	0.28	14.32	926	3.54	3.78	0.23	20.70	1,129

Notes: summary statistics weighted by the longitudinal weight.

Table 3 displays summary statistics of all variables used in the analyses disaggregated by parental SES, adjusted by the longitudinal weight and restricted to their corresponding main analytic samples.

I additionally explored the characteristics of the main subsamples by testing for any systematic differences in terms of socio-demographic and ability composition that could bias comparisons and inference between them. I estimated logistic models to predict the probability of missing data for each subsample as a function of parental SES, gender, migration background, and cognitive and non-cognitive skills, and did not find systematic differences across the subsamples.

4.5. Empirical Strategy

To predict the probability of being enrolled in the academic track by wave 7, I estimate school FE[81] linear probability models (LPM). All models are adjusted by IPW to minimise selective attrition bias, and standard errors are clustered at the school-level to allow for unobserved correlation across students attending the same school. To rule out the possibility of nonlinearities and out-of-bound predictions (around 8%) in LPM, I additionally carried out logistic specifications and operationalised skills in tertiles, obtaining highly consistent results with LPM. I comment on this robustness check throughout the findings' section and additional analyses are shown in the Appendix.

School-FE allow me to control for sources of unobserved confounding that may affect both students' skills and tracking decisions. Namely, school-FE account for all time-constant characteristics of schools and students within the school, such as families' residential and school choice, neighbourhood, school characteristics (e.g., resources, teachers' quality and didactics), and school-specific ability and grading distributions—*grading on a relative curve* according to average classroom performance (Calsamiglia and Loviglio 2017). Around 300 schools with an average of 11 students and 2 classrooms/teachers per school are analysed, depending on model specification and analytic sample. Most variation is explained within schools at around 89% (OECD 2004).

The logic of the empirical analyses is as follows. First, I estimate main models on track attendance to test (1) the direct effect of parental SES, net of skills; (2) the *compensatory*

[81] I also estimated naïve OLS models as a robustness check that are not shown in the chapter, but they yield highly consistent results with School-FE LPM models (results available upon request). This means that between-school confounding may not be a major issue.

advantage hypothesis; and (3) the *skill substitution hypothesis*. Second, I estimate similar models to test whether the research hypotheses on compensation and skill substitution are explained by mechanisms on (1) teachers' bias in GPA and track recommendations, and (2) parental aspirations and expectations.

For the main models on track attendance, six equations are estimated. In equation (1), *i* subscript stands for students within schools, and *j* for schools. y_{ij} is the track decision or type of school currently attended; α represents the intercept or grand mean probability of being enrolled in the academic track across schools. SES_{ij} stands for parental SES and the parameter β_1 accounts for its total effect on track attendance, net of migration status[82]; $\boldsymbol{Z}_{ij}$ is a vector of control variables at the student level (gender and migration background in equations (1-6); and basic cognitive skills in equations (2-6)); δ_j is the school-FE with around 300 dummies, and e_{ij} stands for a student-specific error term.

$$y_{ij} = \alpha + \beta_1 SES_{ij} + \boldsymbol{Z}_{ij} + \delta_j + e_{ij} \quad (1)$$

In equation (2), the following terms are added to equation (1): NCS_{ij} accounts for non-cognitive skills; and CS_{ij} captures domain-specific cognitive skills (competencies). Thus, equation (2) estimates the effect of parental SES on track decisions net of skills (as known as secondary effects) through the parameter β_1, and vice versa.

$$y_{ij} = \alpha + \beta_1 SES_{ij} + \beta_2 NCS_{ij} + \beta_3 CS_{ij} + \boldsymbol{Z}_{ij} + \delta_j + e_{ij} \quad (2)$$

The directed acyclic graph (DAG) in Figure 1 represents equation (2) and the baseline theoretical causal model of this inquiry.[83] Parental SES is considered as an exogenous and antecedent variable to all the rest, and I aim to estimate its approximate causal effect on school choice. As extensively argued above in *sections 3.1.-3.2.*, the temporal sequence of skill variables in this chain of causality was established according to previous research on skill formation dynamics and child development, and data availability in NEPS. I have empirically validated

[82] To avoid over-control and endogenous selection bias, I do not control for any variable that may mediate the total effect of parental SES further than skills, such as family structure or employment status.

[83] To test teachers' bias in track recommendations and parental aspirations/expectations mechanisms, I decided to not assess these variables as mediators between parental SES and track choice to avoid endogenous selection bias. Instead, I use track recommendations and parental aspirations/expectations as outcome variables.

the associations represented by the arrows with semi-partial correlations and School-FE OLS/LPM regressions, as shown at the bottom of Figure 1.

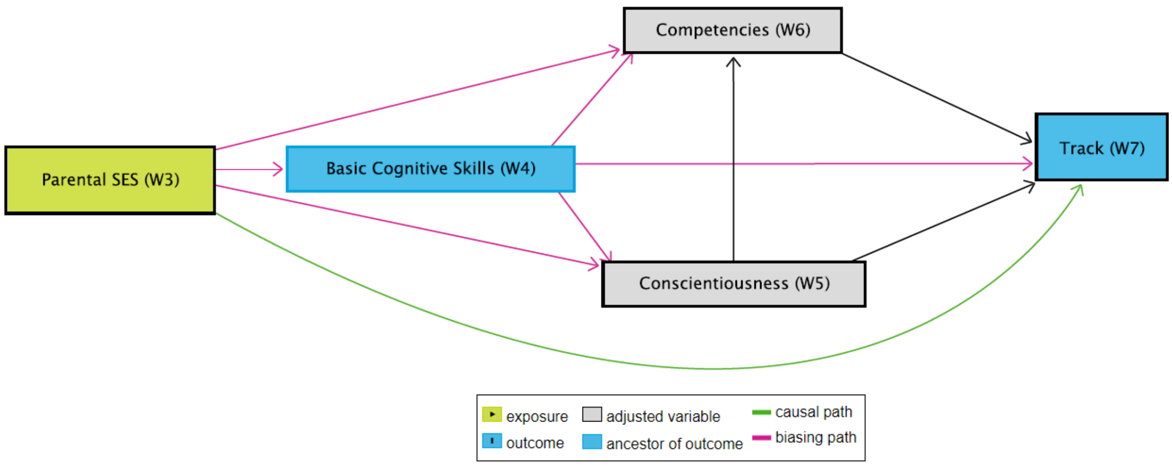


Figure 1. DAG on the basic theoretical causal model

Notes: W=Wave; source: *http://www.dagitty.net/*; Semi-partial correlations (n=2,055): Competencies ⟵ conscientiousness (r = 0.34); basic cognitive skills (r = 0.18); parental ISEI (r = 0.26); Conscientiousness ⟵ basic cognitive skills (r = 0.11); parental ISEI (r = 0.15). Basic cognitive skills ⟵ parental ISEI (r = 0.03).

I make two assumptions and test a third one to identify unbiased, direct effects of parental SES in equation (2). First, I reasonably assume that there are no unobserved confounders considerably affecting both parental SES and track choice. I control for school-FE and migration background, and parental SES is measured as early as elementary school entry[84] (age 6-7 /wave 3) using different proxies (ISEI, income and education).

Second, I assume that there are no unobserved confounders or common causes[85] associated with the mediator variables—cognitive competencies and conscientiousness—and the outcome track choice that could bias estimation through, for instance, endogenous selection bias. Even though this assumption is untestable by definition, I carried out *Seemingly Unrelated Regressions* (SUR) to predict competencies or conscientiousness (first equation) and track choice (second equation) in a system of linear equations to allow for contemporaneous cross-equation error correlation, and controlling for the same predictors (school-FE, sex, migration

[84] As explained above, I consider parental ISEI as time-constant. A reasonable assumption given its high intra-class correlation of 0.73 and low within-individual standard deviation of 6.3. It should be noted, though, that the direct effect of parental SES over and above skills measured during elementary education could be slightly overestimated if the SES-gap in skills is mainly shaped during pre-school age (Passaretta et al. 2020).

[85] For instance, *anticipatory decisions* (or other unobserved choice-based mechanisms) of dropping-out or enrolling in the vocational training tracks may be a potential confounder (Morgan 2012). However, as argued above, at ages 6-10 children have low agency and understanding of their career or educational goals for anticipated decisions to be a relevant confounder.

background, basic cognitive skills, competencies (only in the first equation) and/or conscientiousness (in the second equation, only to predict competencies). *Seemingly Unrelated Regressions* analysis shows that the error terms or residuals of first and second equations are orthogonal. This result holds even when using cognitive and non-cognitive skills' composites as measures of skills and mediators (weighted averages across grades 1-to-4), which potentially neutralise the threat of time-varying confounding.

Third, in equations (3-6) I test for the assumption of no interactions between parental SES and mediators to estimate unbiased direct effects. As shown in Table 3, and as expected, there are significant interactions between cognitive competencies and SES (and between competencies and conscientiousness). Hence, I carried out a causal mediation analysis using parametric regression models to estimate controlled direct effects of parental SES in equation (2). These models thus allow for *treatment* (parental SES)-mediator interactions in the outcome regression model using counterfactual definitions of direct and indirect effects.

In the next step, equation (3) estimates the interaction term between parental SES and non-cognitive skills to test the *compensatory advantage hypothesis*, holding cognitive skills constant. A negative parameter β_4 is expected, so that non-cognitive skills are less predictive of transiting to the academic track at high-SES, and vice versa.

$$y_{ij} = \alpha + \beta_1 SES_{ij} + \beta_2 NCS_{ij} + \beta_3 CS_{ij} + \beta_4 SES_{ij} * NCS_{ij} + \boldsymbol{Z}_{ij} + \delta_j + e_{ij} \quad (3)$$

By the same token, equation (4) estimates the interaction term between parental SES and cognitive skills to test the *compensatory advantage hypothesis*, remaining non-cognitive skills constant. A negative parameter β_4 is expected, so that cognitive skills are less predictive of transiting to the academic track at high-SES, and vice versa.

$$y_{ij} = \alpha + \beta_1 SES_{ij} + \beta_2 NCS_{ij} + \beta_3 CS_{ij} + \beta_4 SES_{ij} * CS_{ij} + \boldsymbol{Z}_{ij} + \delta_j + e_{ij} \quad (4)$$

Equation (5) tests the *skill substitution hypothesis* by adding the interaction term between non-cognitive and cognitive skills, parental SES remaining constant. A negative slope of the parameter β_4 would be evidence on *skill substitution.*

$$y_{ij} = \alpha + \beta_1 SES_{ij} + \beta_2 NCS_{ij} + \beta_3 CS_{ij} + \beta_4 NCS_{ij} * CS_{ij} + \boldsymbol{Z}_{ij} + \delta_j + e_{ij} \quad (5)$$

Finally, equation (6) accounts for the fact that non-cognitive and cognitive skills are not independent by estimating the three-way interaction term between parental SES, non-cognitive and cognitive skills. Therefore, equation (6) formally estimates and tests whether *skill substitution* varies by parental SES, and if it is more prevalent (or only found) at high-SES as represented by a negative slope of parameter β_7. Equations 1-to-6 are estimated to predict track attendance as models 1A-to-6A in Table 4.

$$y_{ij} = \alpha + \beta_1 SES_{ij} + \beta_2 NCS_{ij} + \beta_3 CS_{ij} + \beta_4 SES_{ij}{*}NCS_{ij} + \beta_5 SES_{ij}{*}CS_{ij} + \beta_6 NCS_{ij}{*}CS_{ij} + \beta_7 SES_{ij}{*}NCS_{ij}{*}CS_{ij} + \boldsymbol{Z}_{ij} + \delta_j + e_{ij} \quad (6)$$

To test mechanisms accounting for the compensatory and skill substitution hypotheses, I estimate six school-FE LPM to predict the probability of (i) academic track recommendations, (ii) parental aspirations, and (iii) parental expectations. Equations (2) and (6) are estimated with y_{ij} standing for the corresponding outcome and including the same predictors. However, for the outcome on track recommendations, $\boldsymbol{Z}_{ij}$ includes two additional controls: (1) parental aspirations at school entry (grade 1 /wave 3) to approximate teachers' bias and avoid endogeneity with teachers' expectations (self-fulfilling prophecies)[86]; and (2) type of recommendation issued (consultation and/or formal recommendation). Models 2C, 2D and 2E in Table 6 estimate equation (2) to provide information on SES-gaps net of skills in track recommendations and parental aspirations and expectations, respectively. Moreover, as in equation (6), Models 6C, 6D and 6E in Table 6 test whether *skill substitution* varies by parental SES in predicting academic track recommendations, parental aspirations, and expectations, respectively, with a three-way interaction.

Finally, to test the mechanism on *teachers' bias* by parental SES in assigning grades, I estimate three school-FE OLS models with the same parameters as in equation (2), equation (3), and equation (4) above. However, here y_{ij} stands for the average annual GPA in mathematics and German and α for the grand mean GPA across schools. In model 2B in Table 5, an estimation of teachers' bias in grading is provided by controlling for cognitive and non-cognitive skills. In models 3B and 4B, I test whether teachers' bias is concentrated among students with low non-cognitive skills, or low cognitive skills, respectively.

[86] Given that in wave 6 (2016) most parents (96%) are interviewed between March and May and, in some Federal States track recommendations are issued at the end of the first semester (February), parental educational aspirations and expectations measured at grade 4 may reflect this constraint. For this reason, I use aspirations at school entry (wave 3 at grade 1) as a control in the models predicting track recommendations.

For moderation analyses, I tested assumptions on common support and linearity of the moderator and carried out nonparametric specifications of the interactions (see Appendix Figures A.1.–A.5.). Furthermore, in addition to the main analyses, throughout the chapter I will comment on several robustness checks on confounding, alternative specifications of parental SES (see Appendix Tables A.2.–A.3.) and skills (see Appendix A.1. on measurement error and Tables A.1., A.4., A.5.), among other sensitivity checks on selective attrition bias and selection bias.

5. Findings

Table 4 displays the main results of the multivariable analysis on track choice. Model 1A shows the average total effect of parental SES on track choice—net of migration background and time-constant school characteristics. Students from high-SES families are, on average, 23%[87] more likely to attend the academic track by grade 5/wave 7 than low-SES families (**β** = 0.227 [SE = 0.0344]; *p-value* at 1%). When controlling for cognitive and non-cognitive skills in Model 2A, so accounting for SES-gaps in skills' distributions, the advantage of high-SES families in attending the academic track declines by about half (**β** = 0.128 [SE = 0.0287]; *p-value* at 1%). Still, high-SES families at the same level of skills as low-SES families are 13% more likely to attend the academic track. More precisely, the average direct effect[88] of parental SES accounts for 56.4% of its average total effect (0.128/0.227=0.564), while the average indirect effect through cognitive and non-cognitive skills accounts for the remaining 43.6% (1-0.564=0.436).[89]

Models 3A–4A in Table 4 test the *compensatory advantage hypothesis.* Model 3A rejects that high-SES families compensate for low conscientiousness, as the interaction term is close to a null-effect and non-statistically significant (**β** = -0.0120 [SE = 0.0239]). However, Model 4A does give support to the *compensatory hypothesis* given the negative and significant interaction term between domain-specific cognitive skills (competencies) and parental SES (**β** = -0.0715 [SE = 0.0268]; *p-value* at 1%) by which the marginal effect of cognitive skills is lower among high-SES students. Put it in other words, the SES-gap in transition rates to the academic track is considerably larger among students with low cognitive competencies in mathematics and

[87] 22% for high household income and 27% for high parental education—see Tables A.2.–A.3.
[88] Under the assumptions of (1) no mediator(s)-outcome confounding, and (2) no treatment-mediator(s) interaction (Acharya et al. 2016). Causal mediation analysis allowing for treatment-mediator interaction shows robust results. See also Models M3A–M6A allowing for treatment-mediator interactions.
[89] Models M1F-M2F in Appendix Table A.1. using skills' composite measure yield equivalent results.

German (AME = 0.205 [SE = 0.044]; *p-value* = 0.000) than at medium (AME = 0.134 [SE = 0.029]; *p-value* = 0.000) or high (AME = 0.062 [SE = 0.034]; *p-value* = 0.069) competencies, as estimated with average marginal effects (AME) and linear combinations of coefficients.

Table 4. LMP of academic track

	M1A School-FE	**M2A** School-FE	**M3A** School-FE	**M4A** School-FE	**M5A** School-FE	**M6A** School-FE
Female (male)	0.00808	-0.0656**	-0.0653**	-0.0641**	-0.0688***	-0.0661***
	(0.0319)	(0.0257)	(0.0258)	(0.0256)	(0.0254)	(0.0250)
1st Gen. migrant (non-migrant)	0.250**	0.239**	0.237**	0.239**	0.233**	0.249**
	(0.122)	(0.101)	(0.101)	(0.101)	(0.0984)	(0.0994)
2st Gen. migrant	0.0480	0.0496	0.0496	0.0487	0.0500	0.0512
	(0.0393)	(0.0373)	(0.0373)	(0.0377)	(0.0371)	(0.0375)
High parental ISEI (Low)	0.227***	0.128***	0.129***	0.134***	0.128***	0.149***
	(0.0344)	(0.0287)	(0.0291)	(0.0289)	(0.0289)	(0.0315)
Z - Basic cognitive skills (grade 2)		0.0434***	0.0433***	0.0432***	0.0441***	0.0439***
		(0.0136)	(0.0136)	(0.0134)	(0.0137)	(0.0134)
Z - Conscientiousness (grade 3)		0.0771***	0.0832***	0.0775***	0.0808***	0.0693***
		(0.0158)	(0.0200)	(0.0156)	(0.0160)	(0.0202)
Z - Competencies in maths/German (grade 4)		0.219***	0.220***	0.257***	0.219***	0.261***
		(0.0165)	(0.0165)	(0.0214)	(0.0168)	(0.0226)
Z - Conscientiousness X parental ISEI			-0.0120			0.0310
			(0.0239)			(0.0269)
Z - Competencies X parental ISEI				-0.0715***		-0.0732**
				(0.0268)		(0.0308)
Z - Conscientiousness X Z - competencies					-0.0291**	-0.000248
					(0.0116)	(0.0154)
Z - Consc. X Z - competencies X parental ISEI						-0.0535**
						(0.0249)
constant	0.442***	0.582***	0.582***	0.586***	0.598***	0.588***
	(0.0519)	(0.0420)	(0.0421)	(0.0416)	(0.0420)	(0.0420)
Observations	2,055	2,055	2,055	2,055	2,055	2,055
Schools	292	292	292	292	292	292
R-squared	0.317	0.524	0.524	0.528	0.527	0.532

Notes: Robust standard errors in parentheses; *** p<0.01; ** p<0.05; * p<0.1

Note that this compensatory pattern cannot be explained away by SES-differences in average levels of conscientiousness, so that cognitively weak high-SES students were more conscientious than their low-SES counterparts or ranked higher in other unobserved skills associated with GPA and tracking. To rule out this possibility, Model 4A in Table 3 controls for conscientiousness, and Model 4F in Appendix table A.1. using skills' composites displays equivalent results. Furthermore, as we will see below, Model 6F includes an interaction between competencies and conscientiousness to account for their joint distribution by SES, and the pattern of high-SES compensation for low cognitive competencies holds.

As hypothesised, the observed pattern of *compensatory advantage* for low cognitive competencies among high-SES students can be further unpacked by testing for *skill substitution*. If that were true, we should observe that wealthy but cognitively weak kids at least need to show to their parents and teachers that they can make it through the challenging academic track by being highly perseverant.

Firstly, Model 5A in Table 4 includes an interaction term between cognitive competencies and conscientiousness to test for *skill substitution*, parental SES remaining constant. Indeed, given the negative and statistically-significant interaction term between these cognitive and non-cognitive skills (**β** = -0.0291 [SE = 0.0116]; *p-value* at 5%), it can be claimed that *skill substitution* is at play in predicting tracking outcomes.

Secondly, in order to tease out if *skill substitution* varies by parental SES, Model 6A in Table 4 illustrates that there is a negative and statistically-significant three-way interaction between parental SES, conscientiousness and competencies (**β** = -0.0535 [SE= 0.0250]; *p-value* = 0.033). This means that *skill substitution* is more prevalent (and only found) among high-SES students (**β** = -0.0537 [SE = 0.0192]; *p-value* = 0.006) than low-SES students (**β** = -0.0002 [SE = 0.0154]; *p-value* = 0.987). As shown more clearly in Figure 2 with predicted probabilities estimations, this negative three-way interaction largely suggests that, for high-SES students, conscientiousness pays off more at low cognitive competencies, while at high competencies, conscientiousness is virtually not predictive of school choice. By the same token, cognitive competencies are less predictive at medium/high levels of conscientiousness. This means that high-SES students lacking cognitive or non-cognitive skills are able to substitute/compensate with non-cognitive or cognitive skills, respectively.

On the contrary, as shown in Table 4 Model 6A, for low-SES students the slope of conscientiousness does not vary by cognitive competencies' levels (**β** = -0.0002; [SE = 0.0154]), and vice versa. To put it bluntly, high-SES students that are *not smart but* (are perceived by their teachers as) *hardworking*, tend to get considerably larger access to the academic track than low-SES students with the same combination of skills, where the SES-gap in academic track attendance is the largest (AME = 0.301; [SE = 0.068]; *p-value* = 0.000).

Nonparametric specifications categorising skills in tertiles, as shown in Appendix Figure A.1., give additional support to the *skill substitution* hypothesis. Furthermore, I run a robustness check using GPA as the main measure of cognitive skills and results are consistent with models using cognitive competencies (results available upon request).

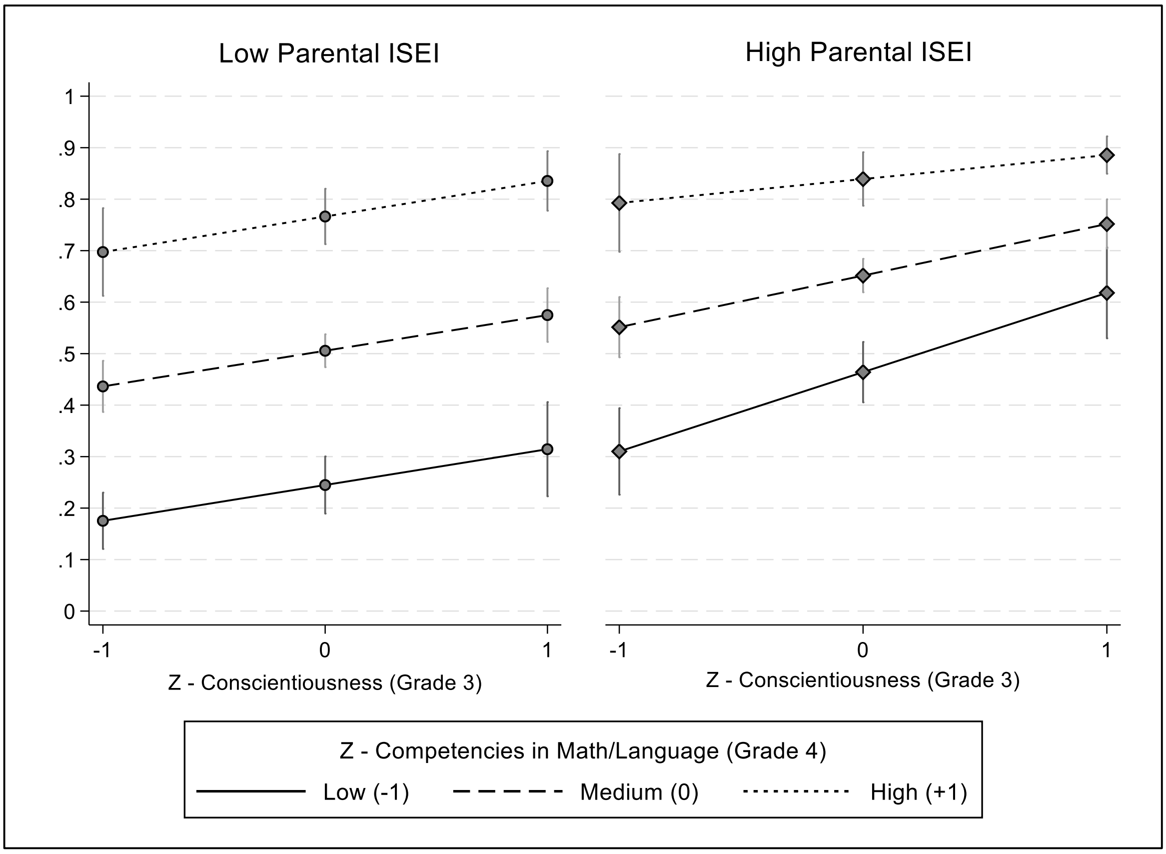


Figure 2. Predicted probabilities of academic track by skills and parental ISEI (95% C.I.)
Notes: Model 6A / Table 4

5.1. Mechanisms

To shed light on the mechanisms underlying the observed patterns of *compensatory advantage* for low cognitive skills and *skill substitution* among high-SES families, in the remainder of this chapter I will focus on the role of teachers' bias in grading and track recommendations, and parental educational aspirations and expectations for their children.

5.1.1. Teachers' Bias

Teachers' bias in *transforming* students' skills into GPA is the first candidate mechanism accounting for the *compensatory effect* in school choice. Table 5 displays Models M2B–M4B testing for teachers' bias in assigning mean GPA in mathematics and German at grade 4.

Model 2B indeed shows that, controlling for cognitive competencies in mathematics and German and conscientiousness, students coming from high-SES families obtain, on average, 0.09 higher GPA than low-SES families (SE= 0.0501; *p-value* at 10%) in a 1 (highest grade)-to-6 (lowest grade) scale with SD=0.79, which means around a 10%- point and 10%-SD higher grade.

Still, one could maintain that this effect is only statistically significant at 10% and that other (unequally distributed by SES) non-observed skills may explain away this SES-gap (or teacher's bias) in GPA. However, in Model M2I in Appendix Table A.4., I use skills' composite measures accounting for any measurable cognitive and non-cognitive skills available to find a similar SES-gap in GPA at 0.116 or 12% a point (SE = 0.043; *p-value*=0.000).

Table 5. OLS regressions on mean GPA in mathematics/German in grade 4

	M2B	**M3B**	**M4B**
	School-FE	School-FE	School-FE
Female (male)	0.119***	0.120***	0.114***
	(0.0431)	(0.0434)	(0.0426)
1st Gen. migrant (non-migrant)	0.0557	0.0586	0.0552
	(0.131)	(0.130)	(0.131)
2st Gen. migrant	0.146**	0.144**	0.143**
	(0.0629)	(0.0633)	(0.0639)
High parental ISEI (low)	-0.0909*	-0.154*	-0.222***
	(0.0501)	(0.0925)	(0.0825)
Z - Basic cognitive skills (grade 2)	-0.0989***	-0.0981***	-0.0997***
	(0.0258)	(0.0257)	(0.0259)
q2 - Conscientiousness (q1) (grade 3)	-0.155***	-0.209**	-0.154***
	(0.0532)	(0.0829)	(0.0532)
q3 - Conscientiousness	-0.308***	-0.347***	-0.310***
	(0.0686)	(0.121)	(0.0702)
q2 - Competencies (q1) (grade 4)	-0.497***	-0.496***	-0.580***
	(0.0685)	(0.0677)	(0.0942)
q3 - Competencies	-0.852***	-0.853***	-0.941***
	(0.0732)	(0.0736)	(0.0936)
q2 - Conscientiousness X parental ISEI		0.104	
		(0.114)	
q3 - Conscientiousness X parental ISEI		0.0755	
		(0.150)	
q2 - Competencies X parental ISEI			0.185
			(0.112)
q3 - Competencies X parental ISEI			0.185**
			(0.0941)
Constant	2.577***	2.606***	2.639***
	(0.0753)	(0.0870)	(0.0839)
Observations	1,774	1,774	1,774
Schools	283	283	283
R-squared	0.550	0.550	0.552

Notes: Robust standard errors in parentheses; ***p<0.01; **p<0.05; *p<0.1

As a next step, Models M3B–M4B test whether teachers' bias by parental SES in grading does vary by students' skills. As far as conscientiousness is concerned, Model 3B shows that the SES-gap in GPA is slightly higher among low-conscientious students, but differences with higher tertiles are not statistically significant.[90] Left-hand Figure 3 illustrates this pattern where the y-axis stands for average annual GPA at grade 4, the y-axis expresses tertiles of conscientiousness at grade 3, and the dotted horizontal lines represent the actual mean GPA of students enrolled in the academic (green line) or vocational tracks (red line).

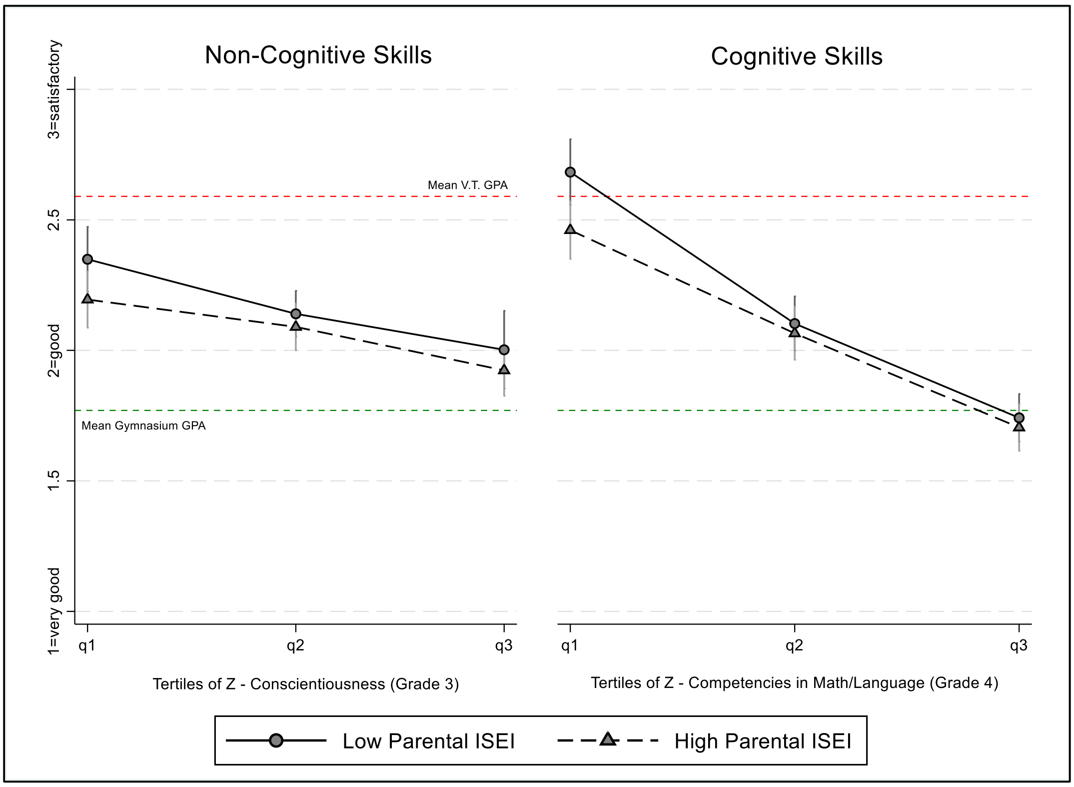


Figure 3. Predicted margins of mean GPA in mathematics/German by skills and parental ISEI (95% C.I.)
Notes: Models 3B-4B – Table 5

Model 4B and right-hand Figure 3 analyse the interaction between competencies and parental SES. As can be seen, teachers' bias by students' SES in grading is only found among students with low cognitive competencies, keeping conscientiousness constant. This SES-gap among low performers is sizeable at 0.222 (SE = 0.082; *p-value*=0.007), the equivalent of a 1/3 SD, and it is found around a critical GPA threshold (2.5) in which some German Federal States

[90] Model M3I in Appendix Table A.4., using a composite measure of non-cognitive skills throughout elementary education, yields a statistically-significant interaction, so that the SES-gap in GPA is concentrated at low levels of non-cognitive skills, being inexistent at the medium and highest tertiles (see Figure A.2.).

apply official guidelines according to the GPA at grade 4 to grant recommendations –only students with a GPA ≤ 2.5 (or 2.0) may be given a recommendation to the academic track. Consequently, I argue that high-SES students with low competencies, independently of their non-cognitive skills as there is no three-way interaction, experience advantages in teachers' evaluations.

Table 6. LPM on teachers' track recommendation and parental aspirations and expectations for the academic track

	Teachers' Track Recommendation		**Parental Aspirations**		**Parental Expectations**	
	M2C	**M6C**	**M2D**	**M6D**	**M2E**	**M6E**
	School-FE		School-FE		School-FE	
Female (male)	-0.0291	-0.0302	-0.0264	-0.0275	-0.0200	-0.0250
	(0.0294)	(0.0282)	(0.0300)	(0.0299)	(0.0287)	(0.0275)
1st Gen. migrant (non-migrant)	0.202**	0.198***	0.210**	0.218**	0.184	0.188*
	(0.0824)	(0.0755)	(0.105)	(0.102)	(0.115)	(0.110)
2st Gen. migrant	0.0326	0.0380	0.0678*	0.0692*	0.0192	0.0183
	(0.0405)	(0.0396)	(0.0388)	(0.0388)	(0.0374)	(0.0370)
High parental ISEI (low)	0.0812**	0.109***	0.0941***	0.108***	0.105***	0.117***
	(0.0350)	(0.0382)	(0.0309)	(0.0338)	(0.0306)	(0.0350)
Z - Basic cognitive skills (grade 2)	0.0504***	0.0487***	0.0118	0.0119	0.0260*	0.0279*
	(0.0182)	(0.0175)	(0.0167)	(0.0163)	(0.0146)	(0.0143)
Z - Conscientiousness (grade 3)	0.105***	0.114***	0.0279*	0.0184	0.0639***	0.0494**
	(0.0188)	(0.0302)	(0.0157)	(0.0248)	(0.0170)	(0.0226)
Z - Competencies in maths/German (grade 4)	0.168***	0.181***	0.141***	0.152***	0.210***	0.247***
	(0.0219)	(0.0358)	(0.0204)	(0.0296)	(0.0210)	(0.0273)
Z - Conscientiousness X parental ISEI		0.00875		0.0283		0.0470
		(0.0350)		(0.0324)		(0.0303)
Z - Competencies X parental ISEI		-0.0173		-0.0165		-0.0654*
		(0.0412)		(0.0377)		(0.0387)
Z - Conscientiousness X Z - competencies		-0.00322		0.00637		-0.0243
		(0.0360)		(0.0282)		(0.0199)
Z - Consc. X Z - comp. X parental ISEI		-0.0745**		-0.0438		-0.0349
		(0.0370)		(0.0342)		(0.0292)
Consultation (recommendation)	0.0228	0.0346				
	(0.0643)	(0.0622)				
Consultation and recommendation	0.0342	0.0346				
	(0.0482)	(0.0491)				
Parental aspirations: acad. track (VT) (grade 1)	0.200***	0.196***				
	(0.0395)	(0.0391)				
Constant	0.455***	0.457***	0.723***	0.722***	0.639***	0.658***
	(0.0676)	(0.0685)	(0.0478)	(0.0484)	(0.0446)	(0.0441)
Observations	1,458	1,458	1,780	1,780	1,780	1,780
Schools	271	271	287	287	287	287
R-squared	0.564	0.573	0.432	0.435	0.510	0.520

Notes: Robust standard errors in parentheses; ***p<0.01; **p<0.05; *p<0.1

Teachers' underassessment of low-SES students at this critical threshold may push downwards parental perceptions of the actual skills and possibilities of their kids to succeed in the academic track. In other words, low-SES parents may be particularly sensitive to the signalling information that low GPA provides. By contrast, high-SES parents may put pressure on teachers to get higher GPA for their low-performing kids given their high educational aspirations/expectations. Nonetheless, I cannot identify whether the observed SES-gap in GPA is due to teachers' explicit and/or implicit bias in their perceptions of students' potential and whether (and how) parents may contribute to shaping it.

Teachers' bias by student's SES in track recommendations may add to their observed bias in grading, proxying for similar underlying mechanisms. As shown in Table 6: Model 2C, on average, high-SES students are 8% more likely (SE = 0.0350; *p-value* at 5%) to obtain a positive track recommendation to the academic track[91] than their low-SES classmates at the same level of skills and parental aspirations at school entry.

Again, as discussed above, it could be argued that high-SES students may rank higher in other unobserved skills, so explaining away any observed SES-gaps in teachers' recommendations. Models 2J and 6J in Appendix Table A.5. rule out this possibility by using skills' composites across elementary education. Results are highly consistent (same coefficient of 0.08) with Models 2C and 6C and give strong support to the argument on teachers' bias by students' SES. Again, I cannot identify whether the SES-gap in track recommendations is driven by teachers' implicit/explicit bias by students' SES or by parental pressure and concurrent aspirations/expectations. Instead, I speculate that both mechanisms are not mutually exclusive and may be reinforcing each other.

Model 6C formally tests whether, as was the case for actual track choice, track recommendations are also driven by *skill substitution.* Indeed, as tested in Table 6: Model 6C and illustrated in Figure 4 below, there is a significant three-way interaction (**β** = -0.0745; [SE = 0.0370]; *p-value* at 5%) by which high-SES students with low cognitive competencies but medium/high conscientiousness experience a boosting likelihood of receiving a positive recommendation to the academic track. On the contrary, for high-performing students coming from advantaged families, it does not matter how conscientious they may be as they have a high probability of getting access to the academic track anyway. For low-SES students,

[91] Controlling for recommendation, high-SES students have 10% more likelihood of making the transition into Gymnasium than low-SES students.

though, conscientiousness and competencies are independent in predicting academic track recommendations.

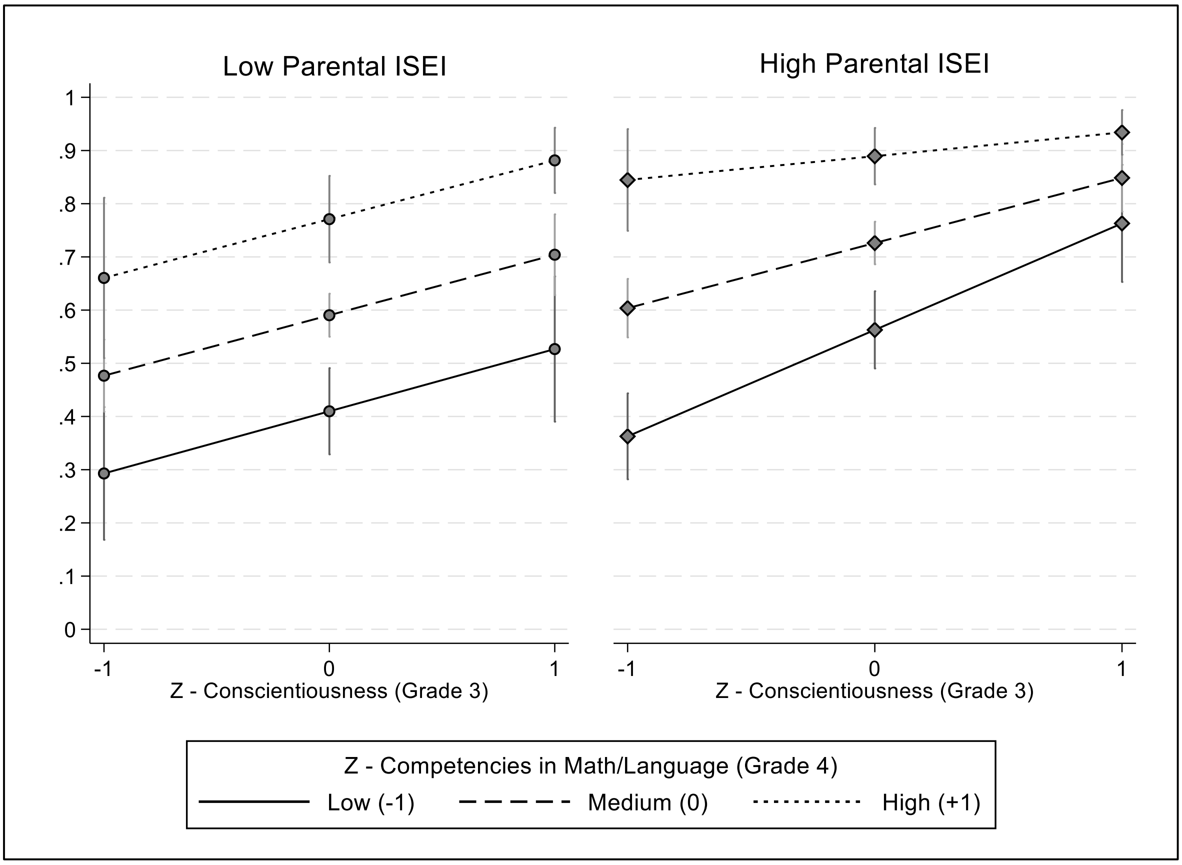


Figure 4. Predicted probabilities of academic track recommendation by skills and parental ISEI (95% C.I.)
Notes: Model 6C-Table 6

To sum up, the negative and statistically-significant three-way interaction between cognitive skills, non-cognitive skills and parental SES suggests that teachers perceive extremely favourably the potential chances of success in the academic track of high-SES students who are cognitively low-performers but strive. However, teachers are substantially more insensitive to how conscientious low-SES students at the bottom of the cognitive distribution may be. Accordingly, the SES-gap or teachers' bias in track recommendations is largest among low-performing (z - competencies at -1) but striving pupils (z - conscientiousness at +1) (AME = 0.201 [SE = 0.089]; p-value = 0.019), as estimated by AME. As illustrated in Appendix Figure A.2., nonparametric specifications measuring skills in tertiles are highly consistent with the finding on *skill substitution*.

5.1.2. Parental Aspirations and Expectations

Parental educational aspirations and expectations for their children is the second key mechanism to be explored in addition to teachers' bias. In Table 6, Models 2D and 6D predict parental aspirations, while Models 2E and 6E predict expectations. As shown in Models 2D and 2E, High-SES parents are about 10% more likely to aspire (**β** = 0.0941 [SE = 0.0309]; *p-value* at 1%) or expect (**β** = 0.108 [SE = 0.0338]; *p-value* at 1%) to the academic track for their kids than low-SES parents, holding skills constant. Similar results are found when using skills' composites in Appendix Table A.5 (M2J–M2K). This finding aligns with previous theoretical predictions and findings on the ambitious educational aspirations and expectations of high-SES families, regardless of ability, likely related to high-status maintenance drives and perceived chances of success.

Models 6D and 6E add the three-way interaction term between cognitive competencies, conscientiousness and parental SES to test for *skill substitution*. For both aspirations (**β** = -0.037; SE = 0.020; *p-value* = 0.063) and expectations (**β** = -0.059; SE = 0.023; *p-value* = 0.009), I find a pattern consistent with *skill substitution* among high-SES families. However, differences with respect to low-SES families are not statistically significant. Note that, in Appendix Table A.5. Models M6J–M6K using skills' composites SES-differences in *skill substitution* are significant at 10%.

Figures 5 and 6 illustrate the predicted probabilities from Models 6D and 6E for parental aspirations and expectations to follow the academic track, respectively. As can be seen, high-SES parents are especially sensitive in their aspirations and expectations when their kids are low-performers but conscientious, being *over*-ambitious. On the contrary, low-SES parents do not considerably change their aspirations or expectations for their low-performing kids if they happen to be highly conscientious. Once again, the largest SES-gap in aspirations and expectations is found among low-performing but striving pupils (aspirations: AME = 0.196 [SE = 0.085]; *p-value* = 0.022; expectations: AME = 0.265 [SE = 0.085]; *p-value* = 0.002), as estimated with AME.

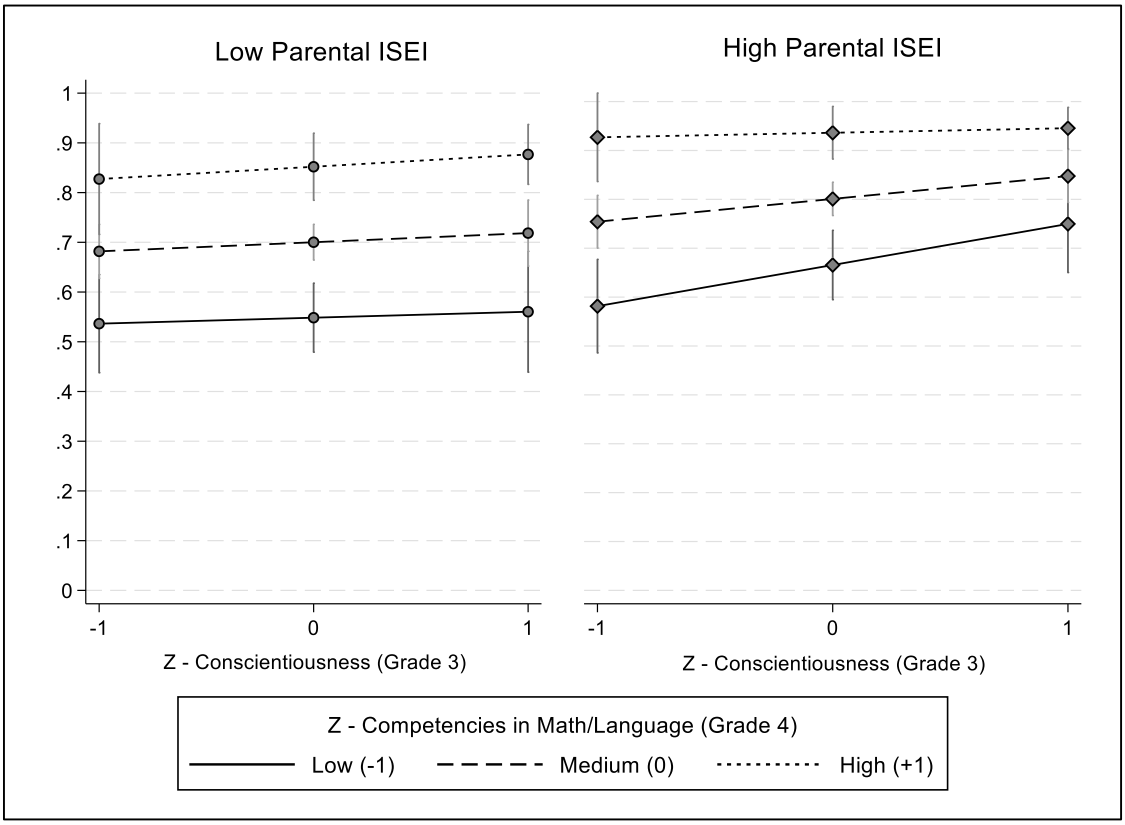


Figure 5. Predicted probabilities of parental academic track aspirations by skills and parental ISEI (95% C.I.)
Notes: Model 6D – Table 6

Nonparametric models categorising skills in tertiles show particularly well this pattern of *skill substitution* by SES in Appendix Figures A.4.–A.5. Moreover, even when parents may be affected by teachers' recommendations when expressing their expectations at the mid-end[92] of grade 4, the similar pattern found for aspirations, which should not be (so) affected by realistic barriers, gives additional support to the findings on expectations.

The SES-gradient in *skill substitution* is especially steep in predicting parental expectations in comparison to aspirations. Parental educational expectations weight in actual barriers and constrains, and I still find very high expectations of high-SES parents among low-performing but striving students in comparison to low-SES parents. Thus, I speculate that this pattern of *over*-ambitious expectations of affluent families for their underperforming kids is highly in line with two central drivers explaining SES-differentials in educational decisions: *perceived chances of success* and *risk aversion to downward mobility* (Barone et al. 2018).

[92] In some Federal States the recommendation comes at the end of the first semester.

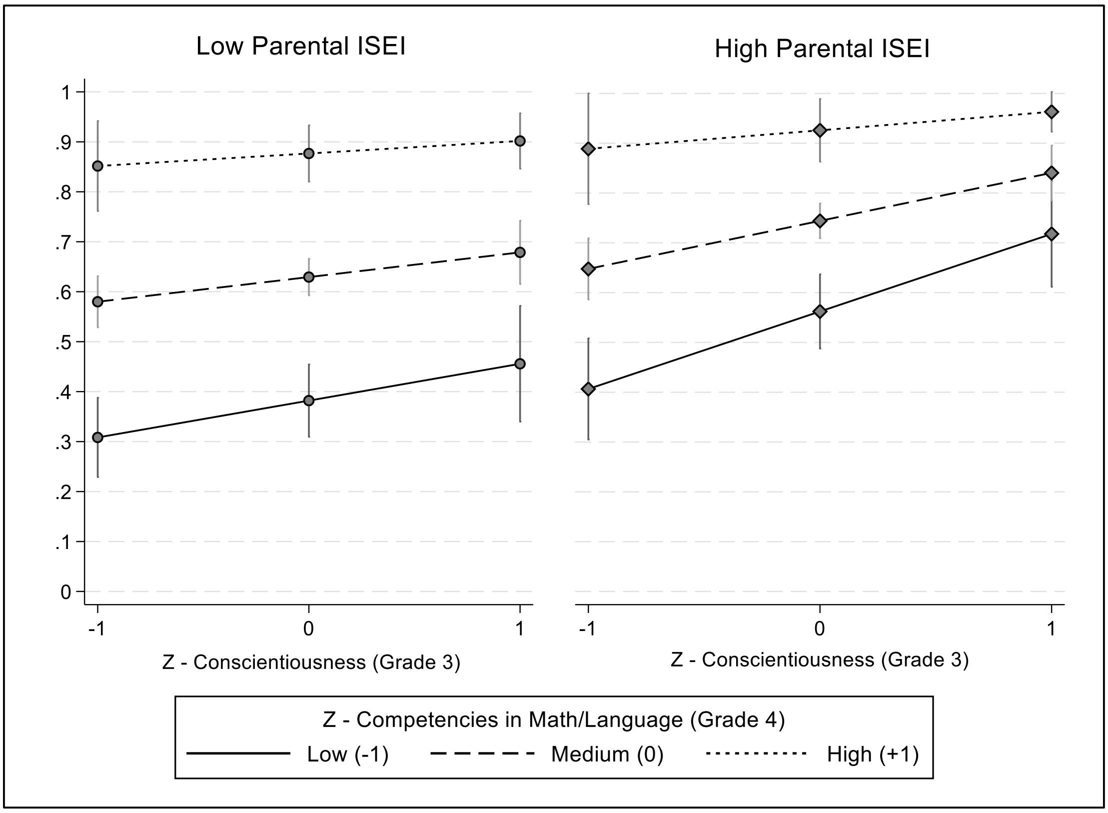


Figure 6. Predicted probabilities of parental academic track expectations by skills and parental ISEI (95% C.I.)
Notes: Model 6E – Table 6

All in all, analyses on mechanisms suggest that teachers' bias in grading and recommendations, and high-SES parents' (over)ambitious expectations for their children go hand-in-hand when it comes to explaining the *compensatory advantage* and *skill substitution* hypotheses and observed empirical findings. However, lacking an experimental design, I cannot disentangle the specific weight of these mechanisms due to the endogeneity of parental and teachers' expectations of students' academic potential. Thus, I argue that both teachers' bias by students' SES and parental expectations may reinforce each other to explain why low-performing (but striving) high-SES students still manage to get disproportionally more access to the academic track than their low-SES classmates.

6. Conclusions

Post-industrial societies are characterised by a zero-sum game of stagnant downward mobility from the upper-echelons and scarce chances of climbing up the ladder among those born at the bottom. The main aim of this chapter was to shed light on this macro process by exploring whether and how low-skilled but advantaged kids tend to avoid downward mobility from early in life.

According to *compensatory theories*, life-course trajectories of pupils from affluent backgrounds are less dependent on prior negative outcomes or disadvantageous traits, such as low cognitive ability, thanks to their resources and status maintenance drives. I tested the *compensatory hypothesis* in the transition into secondary education in Germany, focusing on the roles of cognitive and non-cognitive skills, which are the main predictors of educational and labour market outcomes. These skills are rewarded by school systems and teachers when grading students and are also key indicators of merit for liberal theories of justice to evaluate equality of opportunity in education.

Moreover, skills may interact by being complements or substitutes in predicting learning and educational outcomes, so that low cognitive skills (e.g., mathematics) may be compensated for by high non-cognitive skills (e.g., effort). In this chapter, I argued that *skill substitution* is a potential mechanism accounting for the *compensatory hypothesis* and tested this theory.

Germany represents an ideal context to test *compensatory advantage* and *skill substitutio*n due to its educational system that enforces early tracking into academic or vocational tracks as early as age 10. In this system of early tracking, SES-inequalities in getting a ticket to college by attending academic secondary schools are thought to be mainly driven by SES-gaps in school readiness. Teachers are supposed to objectively grade and recommend tracks as a function of a student's ability and behaviour. Thus, high-SES parents may have less room to compensate if their kids are low performers at the first important educational crossroad for social mobility.

Although previous literature highlighted some potential mechanisms explaining why low-skilled but wealthy students tend to avoid downward mobility, they remain largely under-tested. This chapter provided a three-fold contribution to the literature on intergenerational inequality by testing whether skill substitution, teachers' bias in assessments and parental aspirations are mechanisms underlying the persistency of educational inequalities.

I reported four main findings: (1) high-SES students at the same level of cognitive (e.g., competencies in mathematics/language) and non-cognitive skills (e.g., conscientiousness) as low-SES counterparts are considerably more likely to opt for the academic track; (2) in line with the *compensatory hypothesis*, these socioeconomic inequalities are concentrated among low-performing students (e.g., cognitive competencies in mathematics/language); (3) high-SES students are better able to substitute/compensate for low cognitive skills by high non-cognitive skills in the transition to upper secondary; (4) this heterogeneous pattern of skill substitution by parental SES is likely explained by two complementary mechanisms: (i) teachers' bias by students' SES in grading standards and track recommendations; and (ii) higher educational aspirations/expectations of affluent families. Results were robust to different model and variables' specifications, and several robustness checks on confounding and selective attrition bias.

The analyses testing mechanisms suggested that, on the one hand, teachers perceived high-SES students more favourably by: (i) assigning higher GPA to low-performing (standardised tests in mathematics/language) high-SES students than to their low-SES classmates, keeping non-cognitive skills and IQ constant; and (ii) giving more academic track recommendations to low-performing but striving high-SES students than to low-SES students, even when controlling for previous parental aspirations and IQ. To infer explicit discrimination from these findings is not clear-cut. Nonetheless, if teachers favour characteristics of the student or his/her background that are associated with educational success but not with the student's skills, we can certainly consider these findings as a form of teachers' bias towards low-SES families. Whether these regularities are due to cultural capital, explicit discrimination, or implicit cognitive biases is an open question.

On the other hand, high-SES families express considerably higher educational aspirations and expectations for their low-performing but striving kids than low-SES families. Even when parental educational expectations are supposed to weight in actual barriers and constraints, high-SES parents still express very high expectations for their low ability kids. This finding of *over*-ambitious expectations of advantaged families is consistent with two key mechanisms explaining SES-differentials in educational decisions: *perceived chances of success* and *risk aversion to downward mobility*.

Inspiring popular culture clichés and serious empirical findings claiming that *hard work beats talent*, putting emphasis on the powerful role of perseverance, effort or determination in compensating for low cognitive ability, only seems to work for privileged students in the German educational system—a bottleneck that hinders upward mobility through attending college. The general findings on *compensatory advantage*, *skill substitution*, and *teachers' bias* pose a serious challenge to liberal normative theories of equal opportunity that evaluate merit as the sum of ability plus effort—with teachers being the gatekeepers or evaluators of merit in the school system. Likewise, findings put into question the legitimation of the German system of early-ability tracking based upon selection on meritocratic criteria. As put by Breen and Goldthorpe (2001:82), it seems that "children of less advantaged origins need to show substantially more 'merit' —however understood—than do children from more advantaged origins in order to enter similarly desirable [educational] positions."

This chapter has two main limitations that should be overcome in future research. First, lacking an experimental design, the specific weight of the tested mechanisms due to endogeneity of parental and teachers' expectations of students' academic potential cannot be disentangled. Second, small sample size and attrition prevented heterogeneity analyses by regions to exploit educational legislations, gender or migration background from being carried out. Despite these limitations, this chapter provided solid findings on the interplay between parents, teachers and students' skills in shaping early educational inequalities.

APPENDIX

A.1. Measurement Error

The main measures of cognitive (wave 6; mathematics and language competencies) and non-cognitive skills (wave 5; conscientiousness) studied in this chapter are captured at a single time point with single measures. This strategy induces measurement error bias (e.g., attenuation bias), since teachers tend to evaluate the whole progression of students across elementary education, not just snapshots, and performance in low-stakes standardised tests (e.g., (un)lucky days; non-maximum effort exerted) and teachers' reports of students' behaviour may not capture true ability.[93] To account for measurement error and approximate true ability, I carry out a robustness check in which I rely on several alternative measures of cognitive competencies (13 tests on language, mathematics and scientific literacy) and non-cognitive skills (25 items on motivation, effort, behavioural problems, attention/persistence skills, and personality) taken across grades 1-to-4 (see Table 1 below for details).

Firstly, I apply a latent factor approach (exploratory factor analysis) to test whether cognitive and non-cognitive[94] measures over grades 1-to-4 actually account for two different latent constructs of skills (e.g., unidimensionality). Accordingly, confirmatory factor analysis (principal-components factors) shows that they load strongly on two factors with Eigenvalues of 6.39 and 2.04. As shown in Appendix Figure A.6. and argued above in *section 3.2.*, even though these cognitive and non-cognitive composites are positively correlated (r = 0.45), cognitive and non-cognitive measures capture two differentiated latent factors of skills with satisfactory internal consistency at Cronbach's alpha of 0.85 and 0.83, respectively.

Secondly, to create a composite measure of cognitive competencies across waves independently of age effects, I adapted the strategy by Skopek and Passaretta (2018) by adding the following three steps to the above-explained first step (within-wave standardisation of domains) and second step (within-wave factor analysis of domains): (3) Purging the (wave-specific) mean z-scores from age variation (e.g., maturity effects due to month of birth) by (i) calculating residualised children's age at test date orthogonal to parental SES (e.g., to not underestimate SES effects due to grade retention or month of birth distribution), and (ii)

[93] For some non-cognitive measures, parental and students' reports are available. I additionally built a factor of non-cognitive skills taking parental and students' reports into account and the results hold. However, I decided to only use teachers' reports as the main measure as they better capture children's behaviour in the classroom, which is my focus.
[94] With polychoric correlations as most non-cognitive skills are measured with ordinal Likert scales.

predicting mean z-scores residuals independent from (non-SES related) variation in children's age at test date; (4) standardising the resulting averaged z-scores residuals; and (5) applying factor analysis to estimate the weighted mean z-scores of cognitive competencies across all waves according to the factor loadings of the first (and only retained) factor, with an Eigenvalue of 2.73.

Thirdly, I create a metric composite measure of non-cognitive skills from polytomous ordinal level items (e.g., Likert scales). In order not to violate OLS' and factor analysis' assumptions on linearity, homogeneity and normality (all non-cognitive measures are ordinal and significantly skewed), I applied Item Response Theory (IRT) as a measurement model (Raykov and Marcoulides 2017) that better accounts for the association between the measurement process (Rash models) and the underlying trait to be measured with respect to Classic Test Theory. I run Generalised Partial Credit Models (GPCM) that use mixed logistic regressions to predict the probability of each possible response category to an item (adjacent-categories logits; Rijmen et al. 2003) as a function of the latent trait, the observed responses to the items, and items' parameters - allowing for flexible items' difficulties (e.g., the proportion of persons with a *correct* response in any item defining the construct) and discrimination—item capacity to differentiate subjects (e.g., the slope between the latent trait and the response function). Generalised Partial Credit Models generate sum scores with the properties of a continuous scale (standard normal distribution with mean=0 and SD=1), thus reflecting individual ability on the assumed latent trait and item 'difficulty.' Generalised Partial Credit Models are estimated with MLE, therefore the analysis of items with data missing at random are included. The resulting scores reflect the variability of the pattern of responses and the standard errors account for the level of missing data of each case.

A limitation of this strategy is that composite measures are weighted averages across elementary education and, hence, do not account for growth curve scenarios of skills across grades, as I do not observe all skills at each single grade, and only mathematics and language competencies are suited to study developmental trajectories in NEPS. However, these composite measures of skills across elementary education account for the possibility of reverse causality in the development of cognitive and non-cognitive skills due to cross-productivity dynamics of skill formation.[95]

[95] Alternatively, I estimated models with lagged non-cognitive and cognitive skills measured at previous grades and the results hold.

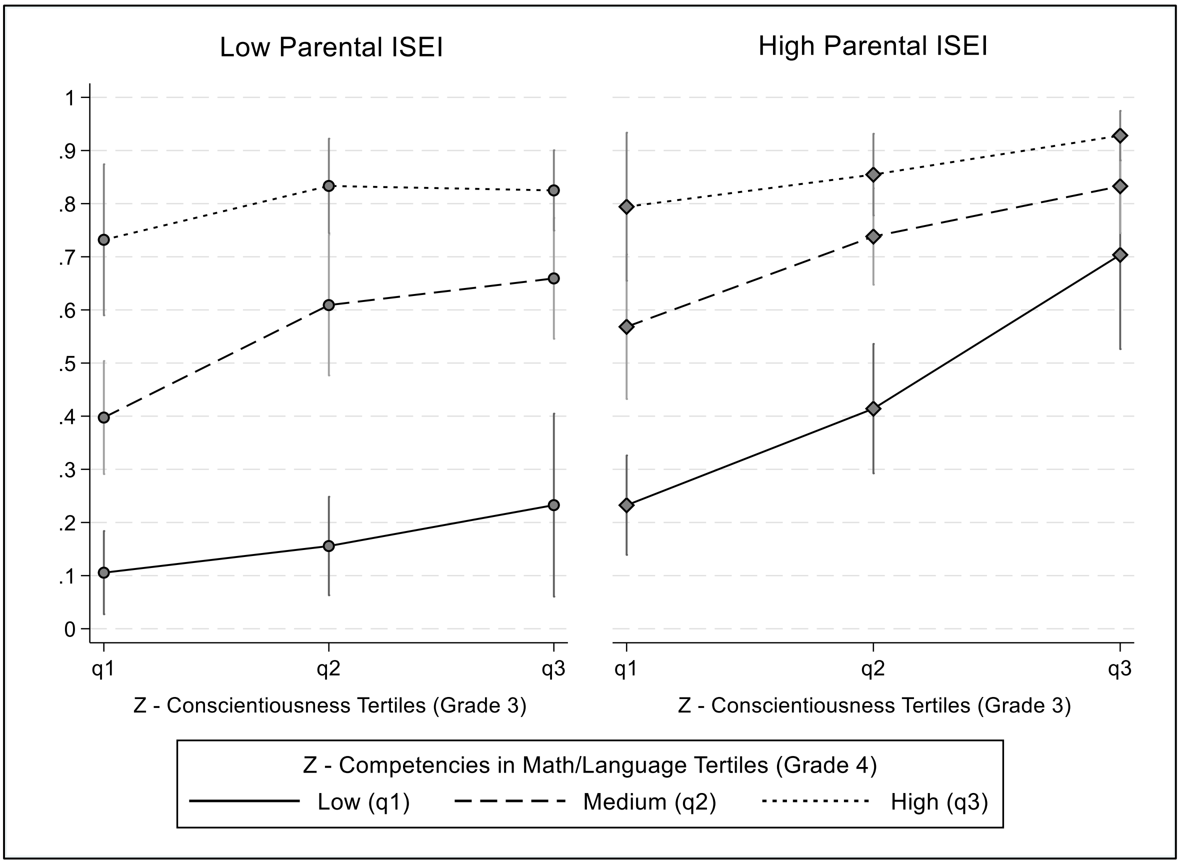


Figure A.1. Linear probability of academic track by skills (tertiles) and parental ISEI (95% C.I.)

Table A.1. LMP of academic track: skills' composites

	M1F	M2F	M3F	M4F	M5F	M6F
	School-FE	School-FE	School-FE	School-FE	School-FE	School-FE
Female (male)	-0.00498	-0.0449*	-0.0430*	-0.0413*	-0.0414*	-0.0385
	(0.0287)	(0.0235)	(0.0236)	(0.0238)	(0.0239)	(0.0241)
1st Gen. migrant (non-migrant)	0.249***	0.286***	0.281***	0.281***	0.275***	0.291***
	(0.0872)	(0.0859)	(0.0845)	(0.0856)	(0.0859)	(0.0868)
2st Gen. migrant	0.0701**	0.121***	0.119***	0.123***	0.119***	0.125***
	(0.0343)	(0.0292)	(0.0293)	(0.0293)	(0.0293)	(0.0289)
High parental ISEI (Low)	0.258***	0.136***	0.142***	0.144***	0.137***	0.171***
	(0.0285)	(0.0235)	(0.0245)	(0.0239)	(0.0236)	(0.0270)
Z - Basic cognitive skills (grade 2)		0.0175	0.0167	0.0169	0.0160	0.0153
		(0.0115)	(0.0114)	(0.0116)	(0.0116)	(0.0115)
Z - Non-cognitive skills' composite (grade 1-4)		0.149***	0.173***	0.151***	0.155***	0.152***
		(0.0144)	(0.0185)	(0.0143)	(0.0148)	(0.0198)
Z – Cognitive competencies' composite (grade 1–4)		0.176***	0.177***	0.223***	0.184***	0.224***
		(0.0130)	(0.0130)	(0.0181)	(0.0136)	(0.0191)
Z - Non-cognitive skills' composite X parental ISEI			-0.0442**			0.0207
			(0.0205)			(0.0232)
Z - Competencies' composite X parental ISEI				-0.0877***		-0.0674***
				(0.0214)		(0.0245)
Z - Non-cognitive skills X Z - competencies					-0.0448***	0.0109
					(0.0120)	(0.0181)
Z - Non-cognitive skills X Z – competencies X parental ISEI						-0.0889***
						(0.0239)
Constant	0.437***	0.511***	0.509***	0.513***	0.523***	0.504***
	(0.0462)	(0.0391)	(0.0393)	(0.0390)	(0.0396)	(0.0400)
Observations	2,680	2,680	2,680	2,680	2,680	2,680
Schools	344	344	344	344	344	344
R-squared	0.313	0.520	0.521	0.526	0.525	0.533

Notes: Robust standard errors in parentheses; ***p<0.01; **p<0.05; *p<0.1

Table A.2. LMP of academic track: household income

	M1G	M2G	M3G	M4G	M5G	M6G
	School-FE	School-FE	School-FE	School-FE	School-FE	School-FE
Female (male)	0.00636	-0.0667**	-0.0670**	-0.0638**	-0.0702***	-0.0662**
	(0.0333)	(0.0268)	(0.0269)	(0.0267)	(0.0265)	(0.0263)
1st Gen. migrant (non-migrant)	0.231*	0.229**	0.229**	0.222**	0.223**	0.214**
	(0.125)	(0.0994)	(0.0999)	(0.100)	(0.0969)	(0.0984)
2st Gen. migrant	0.0249	0.0367	0.0374	0.0327	0.0372	0.0378
	(0.0405)	(0.0372)	(0.0373)	(0.0371)	(0.0368)	(0.0371)
High household income (low)	0.217***	0.126***	0.124***	0.133***	0.128***	0.151***
	(0.0338)	(0.0273)	(0.0278)	(0.0278)	(0.0274)	(0.0296)
Z - Basic cognitive skills (grade 2)		0.0409***	0.0410***	0.0395***	0.0417***	0.0399***
		(0.0142)	(0.0142)	(0.0141)	(0.0141)	(0.0142)
Z - Conscientiousness (grade 3)		0.0787***	0.0706***	0.0792***	0.0827***	0.0587***
		(0.0156)	(0.0195)	(0.0154)	(0.0158)	(0.0189)
Z - Competencies in maths/German (grade 4)		0.221***	0.220***	0.251***	0.220***	0.258***
		(0.0164)	(0.0165)	(0.0212)	(0.0166)	(0.0216)
Z - Conscientiousness X household income			0.0173			0.0597**
			(0.0264)			(0.0289)
Z - Competencies X household income				-0.0615**		-0.0825***
				(0.0282)		(0.0308)
Z - Conscientiousness X Z - competencies					-0.0320***	-0.0117
					(0.0121)	(0.0170)
Z - Conscientiousness X Z - competencies X household income						-0.0466**
						(0.0227)
Constant	0.466***	0.594***	0.594***	0.593***	0.609***	0.599***
	(0.0543)	(0.0446)	(0.0448)	(0.0443)	(0.0445)	(0.0447)
Observations	2,055	2,055	2,055	2,055	2,055	2,055
Schools	292	292	292	292	292	292
R-squared	0.313	0.524	0.524	0.527	0.527	0.533

Notes: Robust standard errors in parentheses; ***p<0.01; **p<0.05; *p<0.1

Table A.3. LMP of academic track: parental education

	M1H	M2H	M3H	M4H	M5H	M6H
	School-FE	School-FE	School-FE	School-FE	School-FE	School-FE
Female (male)	0.00473	-0.0674**	-0.0664**	-0.0622**	-0.0705***	-0.0642**
	(0.0320)	(0.0264)	(0.0264)	(0.0262)	(0.0262)	(0.0260)
1st Gen. migrant (non-migrant)	0.206**	0.214**	0.210**	0.215**	0.209**	0.219**
	(0.103)	(0.0942)	(0.0913)	(0.0922)	(0.0922)	(0.0908)
2st Gen. migrant	0.0577	0.0533	0.0522	0.0557	0.0535	0.0533
	(0.0399)	(0.0380)	(0.0375)	(0.0383)	(0.0378)	(0.0376)
High parental education (low)	0.266***	0.137***	0.143***	0.156***	0.136***	0.174***
	(0.0302)	(0.0253)	(0.0258)	(0.0267)	(0.0254)	(0.0289)
Z - Basic cognitive skills (grade 2)		0.0427***	0.0424***	0.0429***	0.0434***	0.0433***
		(0.0138)	(0.0136)	(0.0136)	(0.0138)	(0.0138)
Z - Conscientiousness (grade 3)		0.0790***	0.0986***	0.0779***	0.0825***	0.0858***
		(0.0155)	(0.0181)	(0.0153)	(0.0158)	(0.0186)
Z - Competencies in maths/German (grade 4)		0.215***	0.214***	0.250***	0.215***	0.247***
		(0.0162)	(0.0161)	(0.0194)	(0.0165)	(0.0199)
Z - Conscientiousness X parental education			-0.0459**			0.00398
			(0.0209)			(0.0251)
Z - Competencies X parental education				-0.0863***		-0.0750***
				(0.0250)		(0.0286)
Z - Conscientiousness X Z - competencies					-0.0279**	0.00189
					(0.0119)	(0.0155)
Z - Consc. X Z - comp. X parental education						-0.0643**
						(0.0248)
Constant	0.457***	0.596***	0.595***	0.591***	0.611***	0.594***
	(0.0509)	(0.0428)	(0.0426)	(0.0424)	(0.0428)	(0.0431)
Observations	2,055	2,055	2,055	2,055	2,055	2,055
Schools	292	292	292	292	292	292
R-squared	0.327	0.524	0.526	0.529	0.527	0.534

Notes: Robust standard errors in parentheses; ***p<0.01; **p<0.05; *p<0.1

Table A.4. OLS regressions on mean GPA in maths/German in grade 4

	M2I	**M3I**	**M4I**
	School-FE	School-FE	School-FE
Female (male)	0.0667*	0.0650*	0.0607*
	(0.0352)	(0.0348)	(0.0350)
1st Gen. migrant (non-migrant)	0.219**	0.223**	0.216**
	(0.110)	(0.107)	(0.108)
2st Gen. migrant	-0.0386	-0.0353	-0.0374
	(0.0533)	(0.0533)	(0.0535)
High parental ISEI (low)	-0.116***	-0.270***	-0.200**
	(0.0434)	(0.0867)	(0.0840)
Z - Basic cognitive skills (grade 2)	-0.0846***	-0.0848***	-0.0859***
	(0.0228)	(0.0229)	(0.0228)
q2 - Non-cognitive skills' composite (q1) (grades 1–4)	-0.340***	-0.411***	-0.336***
	(0.0605)	(0.0806)	(0.0608)
q3 - Non-cognitive skills' composite	-0.584***	-0.706***	-0.584***
	(0.0533)	(0.0820)	(0.0535)
q2 - Cognitive competencies' composite (q1) (grades 1–4)	-0.355***	-0.350***	-0.360***
	(0.0598)	(0.0599)	(0.0859)
q3 - Cognitive competencies' composite	-0.641***	-0.643***	-0.751***
	(0.0624)	(0.0632)	(0.0729)
q2 - Non-cognitive skills' composite X parental ISEI		0.165*	
		(0.0962)	
q3 - Non-cognitive skills' composite X parental ISEI		0.247**	
		(0.106)	
q2 - Cognitive competencies' composite X parental ISEI			0.0274
			(0.110)
q3 - Cognitive competencies' composite X parental ISEI			0.205**
			(0.0866)
Constant	2.774***	2.840***	2.813***
	(0.0835)	(0.0959)	(0.0907)
Observations	2,298	2,298	2,298
Schools	340	340	340
R-squared	0.550	0.553	0.553

Notes: Robust standard errors in parentheses; ***p<0.01; **p<0.05; *p<0.1

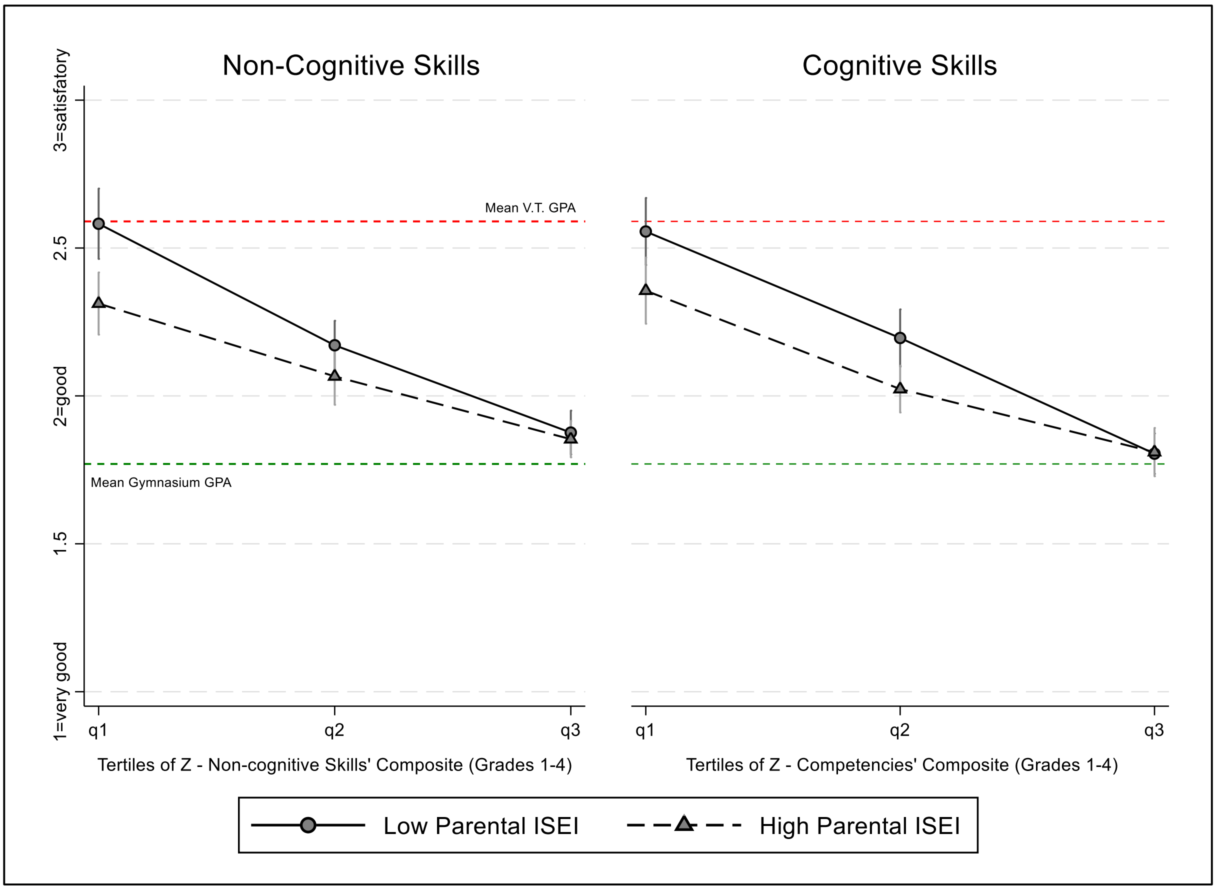


Figure A.2. Predicted margins of mean GPA in maths/German by skills and parental ISEI (95% C.I.)
Notes: Model 2I-3I – Table A.4.

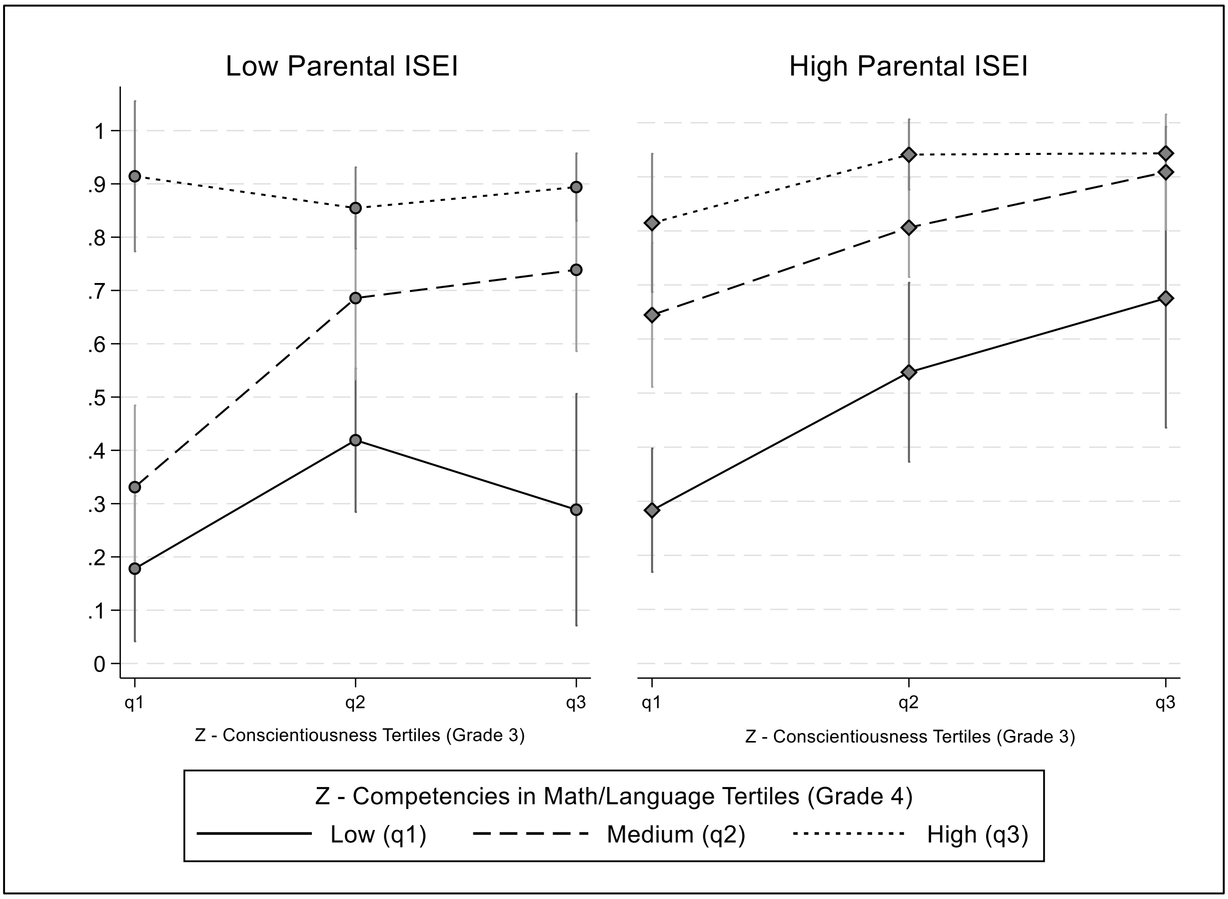


Figure A.3. Linear probability of academic track recommendation by skills (tertiles) and parental ISEI (95% C.I.)

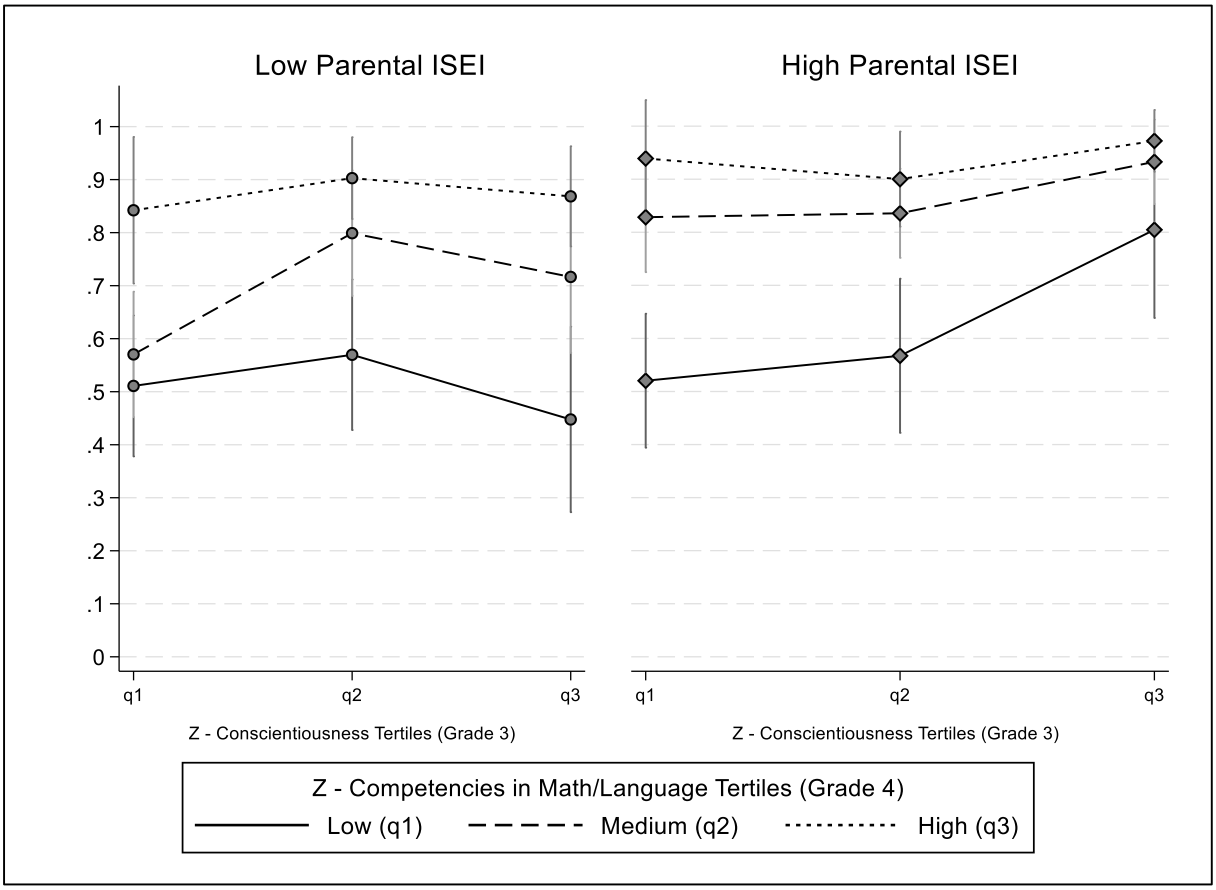


Figure A.4. Linear probability of academic track parental aspirations by skills (tertiles) and parental ISEI (95% C.I.)

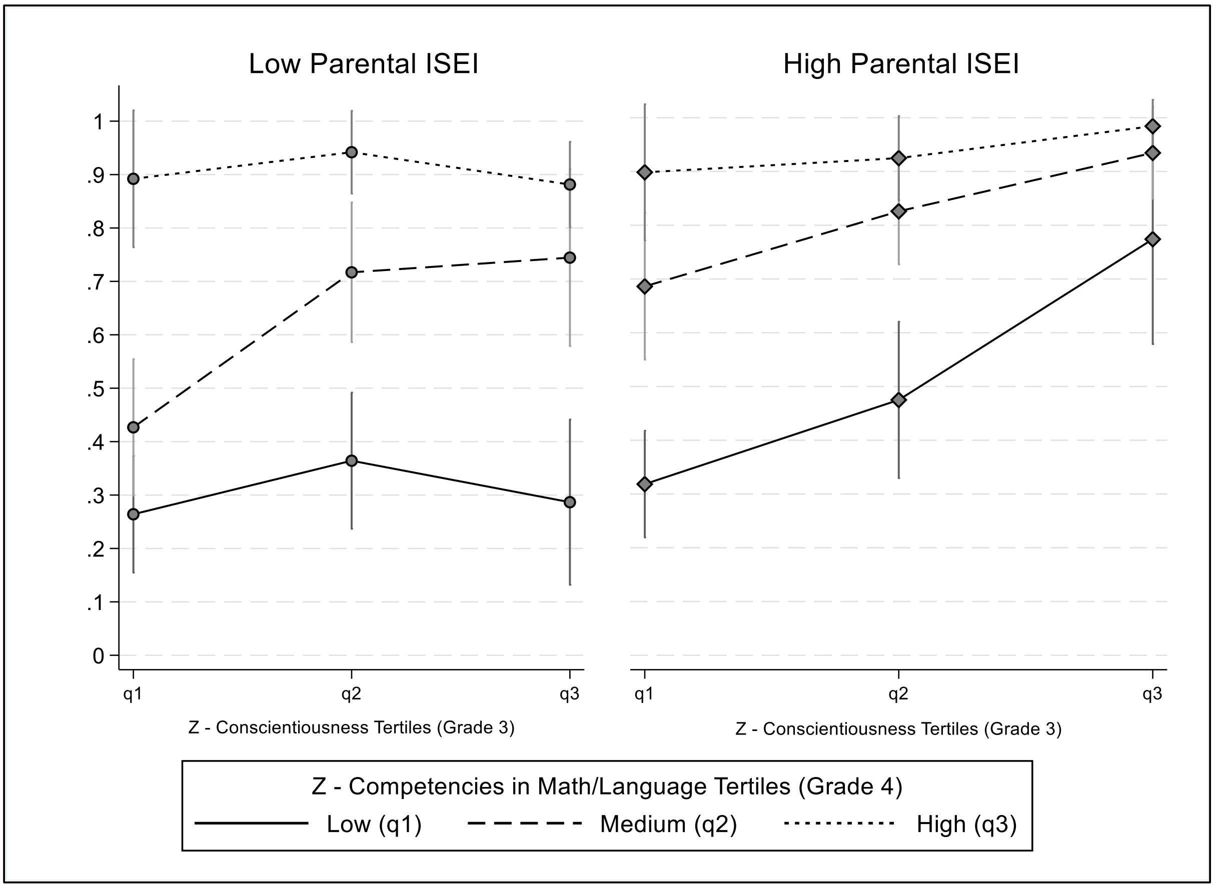


Figure A.5. Linear probability of academic track parental expectations by skills (tertiles) and parental ISEI (95% C.I.)

Table A.5. LPM on teachers' track recommendations and parental aspirations and expectations for the academic track

	Teachers' Track Recommendations		Parental Aspirations		Parental Expectations	
	M2J	**M6J**	**M2K**	**M6K**	**M2L**	**M6L**
	School-FE		School-FE		School-FE	
Female (male)	-0.0283	-0.0252	-0.0213	-0.0189	-0.00744	-0.00176
	(0.0239)	(0.0232)	(0.0237)	(0.0240)	(0.0236)	(0.0240)
1st Gen. migrant (non-migrant)	0.128	0.111	0.205**	0.209**	0.201*	0.194*
	(0.121)	(0.132)	(0.0893)	(0.0892)	(0.104)	(0.103)
2st Gen. migrant	0.0601*	0.0609*	0.111***	0.113***	0.0738**	0.0729**
	(0.0348)	(0.0317)	(0.0305)	(0.0302)	(0.0309)	(0.0311)
High parental ISEI (Low)	0.0802***	0.111***	0.107***	0.128***	0.104***	0.128***
	(0.0294)	(0.0345)	(0.0252)	(0.0268)	(0.0238)	(0.0278)
Z - Basic cognitive skills (grade 2)	0.0197	0.0149	0.00315	0.00148	0.0127	0.00931
	(0.0130)	(0.0129)	(0.0131)	(0.0131)	(0.0118)	(0.0115)
Z - Non-cognitive skills' composite (grades 1–4)	0.144***	0.163***	0.0608***	0.0652***	0.119***	0.130***
	(0.0153)	(0.0274)	(0.0174)	(0.0248)	(0.0166)	(0.0229)
Z - Competencies' composite (grades 1–4)	0.169***	0.202***	0.116***	0.131***	0.174***	0.217***
	(0.0148)	(0.0234)	(0.0143)	(0.0232)	(0.0143)	(0.0208)
Z - Non-cognitive skills' composite X parental ISEI		0.00075		0.00612		0.01000
		(0.0323)		(0.0297)		(0.0274)
Z - Competencies' composite X parental ISEI		-0.0224		-0.0149		-0.0509*
		(0.0313)		(0.0317)		(0.0289)
Z - Non-cognitive skills X Z - competencies		-0.0263		0.0108		-0.0330*
		(0.0255)		(0.0277)		(0.0186)
Z - Non-cognitive Skills X Z - competencies X parental ISEI		-0.0807***		-0.0544*		-0.0519**
		(0.0284)		(0.0312)		(0.0256)
Consultation (recommendation)	0.0252	0.0259				
	(0.0462)	(0.0451)				
Consultation and recommendation	0.0242	0.0127				
	(0.0385)	(0.0378)				
Parental aspirations: acad. track (VT) (grade 1)	0.151***	0.146***				
	(0.0346)	(0.0349)				
Constant	0.465***	0.480***	0.692***	0.685***	0.594***	0.598***
	(0.0531)	(0.0545)	(0.0379)	(0.0394)	(0.0367)	(0.0383)
Observations	1,947	1,947	2,371	2,371	2,371	2,371
Schools	320	320	339	339	339	339
R-squared	0.573	0.592	0.417	0.421	0.506	0.522

Notes: Robust standard errors in parentheses, ***p<0.01; **p<0.05; *p<0.1

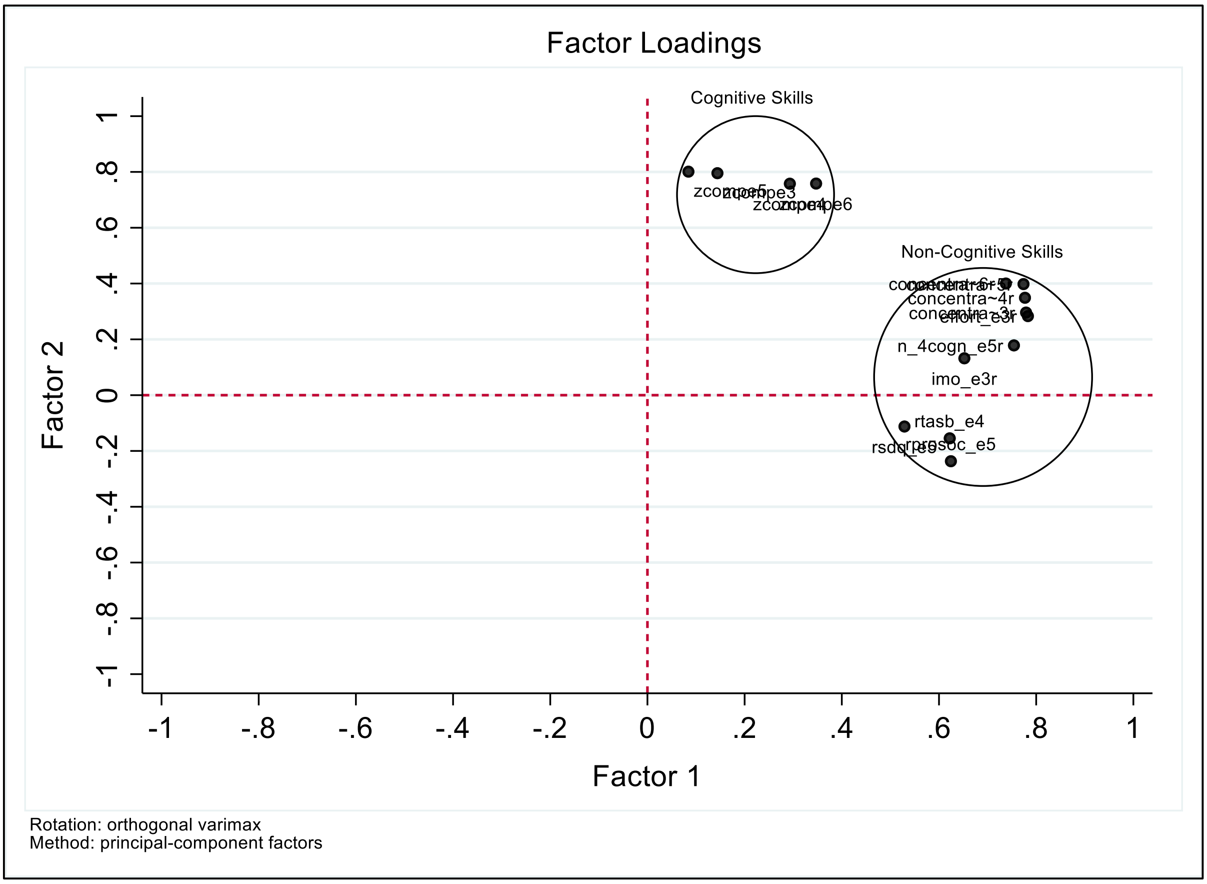


Figure A.6. Factor loadings of cognitive and non-cognitive skills (grades 1–4) from factor analysis

References

Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3), 512–529.

Alesina, A, M Carlana, E La Ferrara and P Pinotti (2018), "Revealing stereotypes: Evidence from immigrants in schools", NBER Working Paper 25333.

Almlund, M., Duckworth, A., Heckman, J.J., Kautz, T., (2011). Personality psychology and economics. In: Hanushek, E.A., Machin, S., Wößmann, L. (Eds.), *Handbook of the Economics of Education*, vol. 4. Elsevier, Amsterdam, pp. 1–181.

Asendorpf, J. B., and Van Aken, M. A. G. (2003). Validity of Big Five Personality Judgments in Childhood: A 9 Year Longitudinal Study. *European Journal of Personality*, *17*(1), 1–17. https://doi.org/10.1002/per.460.

Ashwill, M. A. (1999). *The educational system in Germany: Case study findings.* Washington, DC: National Institute on Student Achievement, Curriculum, and Assessment (ERIC Document Reproduction Service No. ED430906).

Attanasio, O., Blundell, R., Conti, G., and Mason, G. (2020). Inequality in socio-emotional skills: A cross-cohort comparison. *Journal of Public Economics*, 104171. https://doi.org/https://doi.org/10.1016/j.jpubeco.2020.104171

Azzolini, D., Bazoli, N., Lievore, I., Schizzerotto, A., and Vergolini, L. (2019). *Beyond achievement. A comparative look into 15-year-olds' school engagement, effort and perseverance in the European Union.* Luxembourg: Publications Office of the European Union.

Barg, K. (2012). The influence of students' social background and parental involvement on teachers' school track choices: reasons and consequences. *European Sociological Review*, *29*(3), 565–579.

Barone, Carlo, Moris Triventi, Giulia Assirelli, (2018). "Explaining Social Inequalities in Access to University: A Test of Rational Choice Mechanisms in Italy." *European Sociological Review*, 34(5):554–569.

Batruch, A., Autin, F., Bataillard, F., & Butera, F. (in press). School selection and the social class divide: How tracking contributes to the reproduction of inequalities. *Personality and Social Psychology Bulletin.*

Bernardi, F. (2014). Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth. *Sociology of Education*, 87(2): 74–88.

Bernardi, Fabrizio, Cebolla-Boado, Héctor. (2014). "Previous School Results and Social Background: Compensation and Imperfect Information in Educational Transitions." *European Sociological Review*, 30(2):207–14.

Bernardi, Fabrizio, Triventi, Moris. (2018). "Compensatory Advantage in Educational Transitions: Trivial or Substantial? A Simulated Scenario Analysis." *Acta Sociologica*. doi:10.1177/0001699318780950

Bernardi, Fabrizio and Michael Grätz. (2015). "Making Up for an Unlucky Month of Birth in School: Causal Evidence on the Compensatory Advantage of Family Background in England." *Sociological Science*, 2: 235-251.

Blossfeld, H.-P., Rossbach, H.-G, & von Maurice, J. (Eds.) (2011). "Education as a Lifelong Process -- The German National Educational Panel Study (NEPS)." [Special Issue] *Zeitschrift fuer Erziehungswissenschaft*: 14.

Blossfeld et al. (2016a). *Models of Secondary Education and Social Inequality: An International Comparison*. Cheltenham, UK: Edward Elgar.

Blossfeld, H. P., von Maurice, J., Bayer, M., & Skopek, J. (2016b). Methodological issues of longitudinal surveys: The example of the national educational panel study. *Methodological Issues of Longitudinal Surveys: The Example of the National Educational Panel Study*, 1–741.

Boone, S., and Van Houtte, M. (2013). Why are teacher recommendations at the transition from primary to secondary education socially biased? A mixed-methods research. *British Journal of Sociology of Education*, 34(1), 20–38.

Borghans, L., Duckworth, A. L., Heckman, J. J., Weel, B. (2008). The Economics and Psychology of Personality Traits. *The Journal of Human Resources*, *43*, 972–1059.

Borghans, L., Golsteyn, B. H. H., Heckman, J. J., & Humphries, J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences*, *113*(47), 13354 LP – 13359.

Bowles, S. and H. Gintis (1976). *Schooling in Capitalist America*. Basic Books, New York.

Boysen GA and Vogel DL. (2009). Bias in the Classroom: Types, Frequencies, and Responses. Teaching of Psychology, 36(1): 12-17.

Brandt, N. D., Lechner, C. M., Tetzner, J., & Rammstedt, B. (2019). Personality, cognitive ability, and academic performance: Differential associations across school subjects and school tracks. *Journal of Personality*, 0–3.

Breen, R., and Müller, W. (2020). *Education and Intergenerational Social Mobility in Europe and the United States*. Stanford, California: SUP.

Breen, Richard, Goldthorpe, John H. (1997). "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society*, 9(3):275–305.

Breen, R. and Goldthorpe, J. H. (2001). Class, Mobility and Merit: the Experience of two British Birth Cohorts. *European Sociological Review*, 17(2): 81-101.

Briley. D. A. and Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, 24, 1704-1713.

Bub, K. L., Mccartney, K., and Willett, J. B. (2007). Behavior Problem Trajectories and First-Grade Cognitive Ability and Achievement Skills: A Latent Growth Curve Analysis, *Journal of Educational Psychology*, 99(3), 653–670.

Bukodi, Erzsébet, Bourne, Mollie, Betthäuser, Bastian. (2017). "Wastage of Talent? Social Origins, Cognitive Ability and Educational Attainment in Britain." Advances in Life Course Research 34:34–42.

Calero, M. D., Garcia-Martin, M. B., Jimenez, M. I., Kazen, M., & Araque, A. (2007). Self-regulation advantage for high-IQ children: Findings from a research study. *Learning and Individual Differences*, 17, 328–343.

Calsamiglia, Caterina, and Annalisa Loviglio. (2017). "Grading on a curve: When having good peers is not good." Unpublished manuscript.

Conti, G., and Heckman, J. J. (2014). Understanding conscientiousness across the life course: An economic perspective. *Developmental Psychology*, *50*(5), 1451–1459.

Costa, P.T., McCrae, R.R., 1992a. Four ways five factors are basic. *Personality and Individual Differences*, 13(6), 653–665.

Damian, R. I., et al. (2015). Can personality traits and intelligence compensate for background disadvantage? Predicting status attainment in adulthood. *Journal of Personality and Social Psychology*, *109*(3), 473–489.

De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*, *91*(3), 538–552.

Deary, Ian J., Johnson, Wendy. (2010). "Intelligence and Education: Causal Perceptions Drive Analytic Processes and Therefore Conclusions." *International Journal of Epidemiology*, 39(5):1362–69.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007, November 30). Preschool program improves cognitive control. *Science*, 318, 1387–1388.

Duckworth, A. L. and Seligman, M. E. P. (2005) 'Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents', *Psychological Science*, 16(12), pp. 939–944. doi: 10.1111/j.1467-9280.2005.01641.x.

Duckworth, A. L., Quinn, P. D., and Tsukayama, E. (2012). What No Child Left Behind Leaves Behind: The Roles of IQ and Self-Control in Predicting Standardized Achievement Test Scores and Report Card Grades. *Journal of Educational Psychology*, *104*(2), 439–451. https://doi.org/10.1037/a0026280

Duncan, G., and Magnuson, K. (2011). The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems. In: Duncan G., Duncan G., & Murnane R. (Eds.), Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances (pp. 47-70). Russell Sage Foundation.

Eisenberg, N., Duckworth, A. L., Spinrad, T. L., and Valiente, C. (2014). Conscientiousness: Origins in childhood? *Developmental Psychology*, *50*(5), 1331–1349. https://doi.org/10.1037/a0030977

Esser, H. (2016). "The model of ability tracking – Theoretical expectations and empirical findings on how educational systems impact on educational success and inequality." S. 25-42 in: Hans-Peter Blossfeld, Sandra Buchholz, Otto Friedrich, Jan Skopek, Moris Triventi (Eds.) *Models of Secondary Education and Social Inequality. An International Comparison.* Cheltenham: Edward Elgar.

Esping-Andersen, G., and Cimentada, J. (2018). Ability and mobility: The relative influence of skills and social origin on social mobility. *Social Science Research*, *75*(June), 13–31. https://doi.org/10.1016/j.ssresearch.2018.06.003

Farkas, G. (2003). Cognitive Skills and Noncognitive Traits and Behaviors in Stratification Processes. *Annual Review of Sociology*, *29*(1), 541–562.

Fishkin, J. (2014). *Bottlenecks: A New Theory of Equal Opportunity.* Oxford: Oxford University Press.

Fleming, K.A., Heintzelman, S.J., and Bartholow, B.D. (2016). Specifying Associations between Conscientiousness and Executive Functioning: Mental Set Shifting, Not Prepotent Response Inhibition or Working Memory Updating. *Journal of personality, 84 3*, 348-60.

Forster, Andrea G., Herman G van de Werfhorst (2020). Navigating Institutions: Parents' Knowledge of the Educational System and Students' Success in Education, *European Sociological Review*, 36(1): 48–64, Ganzeboom HBG, Treiman DJ. (1996). Internationally comparable measures of occupational status for the 1988 international Standard Classification of Occupations. *Social Science Research*, 25,201-239.

Geven, S., Batruch, A., and van de Werfhorst, H. (2018). *Inequality in Teacher Judgements, Expectations and Track Recommendations: A Review Study*. Amsterdam: Universiteit van Amsterdam.

Goldthorpe, J. (2007). *On Sociology*. Stanford, California: SUP.

Haberkorn, K., Pohl, S. (2013). Cognitive basic skills – Data in the Scientific Use File. Bamberg: University of Bamberg, National Educational Panel Study.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences of the United States of America, 104*(33), 13250-13255. doi:10.1073/pnas.0701362104

Heckman, J. J., and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics, 19*(4), 451–464.

Holm, A., Anders Hjorth-Trolle, and Mads Meier Jæger (2019). "Signals, Educational Decision-Making, and Inequality." *European Sociological Review*, 35(4):447–460

Hsin, A., and Xie, Y. (2016). Hard skills, soft skills: The determinants and consequences of cognitive and socio-behavioral skills. *Social Science Research.*

Jackson, M. (2013). *Determined to succeed? Performance versus choice in educational attainment*, Stanford: Stanford University Press.

Jæger, M.M. (2011). "Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data." *Sociology of Education*, *84(4)*:281–298.

Jæger, M.M., and Møllegaard, S. (2017). "Cultural capital, teacher bias, and educational success: New evidence from monozygotic twins." *Social Science Research*, 65:130–144.

Jürges, H., Schneider, K., (2007). "What Can Go Wrong Will Go Wrong: birthday Effects and Early Tracking in the German School System." CESIfo Working Paper 2055. CESifo. Munich.

Knopik, Valerie S., Neiderhiser, Jenae N., DeFries, John C., Plomin, Robert. (2017). *Behavioral Genetics*. New York: Macmillan Learning.

Krapohl, E., et al. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences, 111*(42), 15273–15278.

Krolak-Schwerdt S., Pit-ten Cate I.M., Hörstermann T. (2018) Teachers' Judgments and Decision-Making: Studies Concerning the Transition from Primary to Secondary Education and Their Implications for Teacher Education. In: Zlatkin-Troitschanskaia O., Toepper M., Pant H., Lautenbach C., Kuhn C. (eds) Assessment of Learning Outcomes in Higher Education. Methodology of Educational Measurement and Assessment. Springer, Cham

Kyllonen, P., and Kell, H. (2018). Ability Tests Measure Personality, Personality Tests Measure Ability: Disentangling Construct and Method in Evaluating the Relationship between Personality and Ability. *Journal of Intelligence*, *6*(3), 32.

Lareau, A. (2015). Cultural Knowledge and Social Inequality. *American Sociological Review*, 80(1): 1–27.

Light, A., and Nencka, P. (2019). Learning and Individual Differences Predicting educational attainment : Does grit compensate for low levels of cognitive ability ? *Learning and Individual Differences*, *70*(January), 142–154.

Liu, A. (2019). Can non-cognitive skills compensate for background disadvantage ? – the moderation of non-cognitive skills on family socioeconomic status and achievement during early childhood and early adolescence. *Social Science Research*, *83*(April), 102306.

Malanchini, M., Engelhardt, L. E., Grotzinger, A. D., Harden, K. P., & Tucker-Drob, E. M. (2018). "Same But Different": Associations Between Multiple Aspects of Self-Regulation, Cognition, and Academic Abilities. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000224

Malanchini, M., Rimfeld, K., Allegrini, A. G., Ritchie, S. J., and Plomin, R. (2020). Cognitive ability and education: How behavioural genetic research has advanced our knowledge and understanding of their association. *Neuroscience & Biobehavioral Reviews*, *111*, 229–245.

Marcenaro-Gutierrez, O., and Vignoles, A. (2015). A comparison of teacher and test-based assessment for Spanish primary and secondary students. *Educational Research*, *57*(1):1–21.

McCrae, R. R. and Mõttus, R. (2019) 'What Personality Scales Measure: A New Psychometrics and Its Implications for Theory and Assessment', *Current Directions in Psychological Science*, 28(4), pp. 415–420.

Meldrum, R. C., Petkovsek, M. A., Boutwell, B. B., & Young, J. T. N. (2017). Reassessing the relationship between general intelligence and self-control in childhood. *Intelligence*, *60*, 1–9. https://doi.org/10.1016/j.intell.2016.10.005

Metcalfe, L. A., Harvey, E. A., and Laws, H. B. (2013). The longitudinal relation between academic/cognitive skills and externalizing behavior problems in preschool children. *Journal of Educational Psychology, 105*(3), 881–894.

Morgan, S. L. (2012) 'Models of College Entry in the United States and the Challenges of Estimating Primary and Secondary Effects', *Sociological Methods & Research*, 41(1), pp. 17–56. doi: 10.1177/0049124112440797.

Morgan, S.L. (1998). Adolescent educational expectations rationalized, fantasized, or both? *Rationality and Society*, 10(2): 131–162.

Murray, A. L., Johnson, W., Mcgue, M., & Iacono, W. G. (2014). How are conscientiousness and cognitive ability related to one another ? A re-examination of the intelligence compensation hypothesis. *Personality and Individual Differences, 70*, 17–22.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). "Intelligence: new findings and theoretical developments." *The American Psychologist*, 67(2), 130–59.

OECD (2004), "How Student Performance Varies between Schools and the Role that Socio-Economic Background Plays in This", in *Learning for Tomorrow's World: First Results from PISA 2003*, OECD Publishing, Paris, https://doi.org/10.1787/9789264006416-5-en.

Passaretta, Giampiero, Jan Skopek, and Thomas van Huizen (2020). "To What Extent Is Social Inequality in School-age Achievement Determined Before and During Schooling? A Longitudinal Analysis in Three European Countries." *SocArXiv*. April 3. doi:10.31235/osf.io/yqt6n.

Polderman, T.J., Gosso, M.F., Posthuma, D., Beijsterveldt, T.V., Heutink, P., Verhulst, F.C., & Boomsma, D.I. (2006). A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta neurologica Belgica, 106 4*, 191-207.

Polderman TJC, Boomsma DI, Bartels M, Verhulst FC, Huizink AC (2010). A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatrica Scandinavica*, 122: 271–284.

Poropat, A. E. (2009). "A meta-analysis of the five-factor model of personality and academic performance." *Psychological Bulletin*, 135:322–338.

Rammstedt, B.; Lechner, C.M.; Danner, D. (2018). "Relationships between personality and cognitive ability: A facet-level analysis." *Journal of Intelligence*. 2018, 6, 28.

Raykov, T., and Marcoulides, G. A. (2017). *Item response theory and modeling using Stata.* College Station, TX: StataPress.

Rijmen, F., Tuerlinckx, F., De Boeck, P., and Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods*, 8, 185–205.

Rikoon, S. H., Brenneman, M., Kim, L., Khorramdel, L., Burrus, J., Roberts, R. D., Roberts, R. D. (2016). Facets of Conscientiousness and their Differential Relationships with Cognitive Ability. *Journal of Research in Personality.*

Roberts, BW.; Jackson, JJ.; Fayard, JV.; Edmonds, G.; Meints, J. Conscientiousness (2009). In: Leary, M.; Hoyle, R., editors. Handbook of individual differences in social behavior. New York, NY: Guilford; p. 369-381.

Rothbart, M. K., and Bates, J. E. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (p. 99–166). John Wiley & Sons Inc.

Shanahan, M., et al. (2014). Personality and the reproduction of social class. *Social Forces*, *93*(1), 209–240. https://doi.org/10.1093/sf/sou050

Skopek, J. and Dronkers, J. (2015), 'Performance in Secondary School in German States - A Longitudinal Three-Level Approach', *Working Paper.*

Skopek, Jan, and Giampiero Passaretta. (2018). “The Social Stratification of Skills from Infancy to Adolescence – Evidence from an Accelerated Longitudinal Design.” *SocArXiv.* October 1. doi:10.31235/osf.io/xkctv.

Smithers, L. G., Sawyer, A. C. P., Chittleborough, C. R., Davies, N. M., Davey Smith, G., and Lynch, J. W. (2018). A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes. *Nature Human Behaviour*, *2*(11), 867–880.

Steinhauer, H. W., Zinn, S., and Aßmann, C. (2016). Weighting Panel Cohorts in Institutional Contexts. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), Methological Issues of Longitudinal Surveys (p. 39-61). Wiesbaden: Springer.

Südkamp, A., Kaiser, J., and Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*(3), 743-762.

Timmermans, A. C., de Boer, H., Amsing, H. T. A., & van der Werf, M. P. C. (2018). Track recommendation bias: Gender, migration background and SES bias over a 20-year period in the Dutch context. *British Educational Research Journal*, 44(5), 847-874

Timmermans, A. C., Kuyper, H., and Werf, G. (2015). Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education? *British Journal of Educational Psychology*, 85(4), 459–478.

Triventi, M. (2019). Are Children of Immigrants Graded Less Generously by their Teachers than Natives, and Why? Evidence from Student Population Data in Italy. *International Migration Review.*

Uka, F., Gunzenhauser, C., Larsen, R. A., and Suchodoletz, A. Von. (2019). Exploring a bidirectional model of executive functions and fluid intelligence across early development. *Intelligence*, *75*(21), 111–121.

Watts, T. W., Duncan, G. J., and Quan, H. (2018). Revisiting the Marshmallow Test: A conceptual replication investigating links between early delay of gratification and later outcomes. *Psychological Science*, 29(7), 1159-1177.

Weinert, S., Asendorpf, J. B., Beelmann, A., Doil, H., Frevert, S., Hasselhorn, M., & Lohaus, A. (2007). Expertise zur Erfassung von psychologischen Personmerkmalen bei Kindern im Altervon fünf Jahren im Rahmen des SOEP

(DIW: Data Documentation 20). Berlin: DIW.

Weinert, S., et al. (2011). *5* Development of competencies across the life span. *ZFE*, *14*(2), 67–86.

Wenz, S. E., and Hoenig, K. (2020). Ethnic and Social Class Discrimination in Education: Experimental Evidence from Germany. *Research in Social Stratification and Mobility*. https://doi.org/10.1016/j.rssm.2019.100461

Zimmermann, T. (2020). Social Influence or Rational Choice? Two Models and Their Contribution to Explaining Class Differentials in Student Educational Aspirations, *European Sociological Review*, 36(1): 65–81.

Chapter V

Conclusions

Carlos J. Gil Hernández

1. Research Question, Contributions and Case Study

This dissertation aimed at answering a central unresolved question in social stratification research: how do wealthy families manage to avoid intergenerational downward mobility from early in life? In a context of stagnant occupational upgrading where the *room at the top* is limited, the only way for the working-classes to experience increasing upward mobility is that upper-class children fall down the social ladder, but they seem to be well-sheltered by their parents resources and drive for status-maintenance.

Attaining a high level of education is surely one of the most effective channels for the upper classes to reproduce their status in post-industrial societies. Indeed, educational inequalities in college enrolment have remained relatively persistent during the last decades, and upper-class families manage to prevent their kids from educational failure even when they happen to be low performers. In Germany, 43% (27%) of low-performing-high-SES kids (low-SES kids) opt for the academic track leading to college, while only 76% (88%) of high-performing- low-SES kids (high-SES kids) make it to academic education. It is not possible to accurately quantify the extent of "deserved" downward mobility that is not happening for high-SES families in Germany but, equalising this proportion could be the first step to improve social mobility rates under the *zero-sum game* of stagnant structural change. This state of affairs has important implications for social mobility and social justice in post-industrial societies and thus deserved further scrutiny in my dissertation.

Particularly, the dissertation focused on Germany with two core research objectives: (1) analysing inequalities by social background in two crucial elements for status-attainment during childhood (age 5–11), i.e., academic ability and educational transitions; and (2) evaluating how advantaged families compensate for negative events, such as LBW and low cognitive ability of their children, for skill formation (cognitive and non-cognitive skills) and educational attainment (transition into academic secondary education).

The dissertation framed these research objectives by drawing from interdisciplinary literature. I mainly tested sociological theories on persistent educational inequality, such as rational action and cultural reproduction theories, and evaluated the implications of the empirical findings for (liberal) theories of social justice and its concept of equal opportunity. I also complemented this framework with contributions from skill formation models in economics; developmental and personality psychology; behavioural genetics models on sources of variation in phenotypic traits; and epidemiological theories on the *foetal origins of disease*. This unusual interdisciplinary approach, along with the evaluation of normative implications for equal opportunity, constituted a substantial contribution to the field of social stratification.

The dissertation provided a two-fold empirical contribution to the literature on intergenerational educational inequality. Firstly, the dissertation studied if BW, an indicator of developmental potential and natural assets in Rawlsian terms, has a heterogeneous effect on children's cognitive and non-cognitive skills by parental SES. This analysis aimed at illustrating how SES-inequalities in academic ability gestate from the prenatal environment in a process of accumulation of small (dis)advantages in which environmental and biological factors interact intricately. The main contribution to the literature lied in exploring two mechanisms accounting for the heterogeneous effect of BW on academic skills by parental SES: (1) relative allocation of investments within families, and the absolute level of investments between families; as well as (2) applying a causal design, drawing from the *Twin Life Study*, that exploited random variation in twins' BW as a natural experiment.

Secondly, the dissertation assessed if, even when controlling for SES-inequalities in students' cognitive (e.g., IQ) and non-cognitive skills (e.g., effort), pupils from advantaged social origins at similar levels of (liberal) academic merit than more disadvantaged schoolmates are more likely to opt for academic tracks bound to college. The German system of education funnels students into academic or vocational pathways as early as age 10. Thus, I considered this system of early tracking as a starting gate where the selection criterion is supposed to be grounded on academic ability at the end of public elementary education when fair life chances are thought to have been already granted. The main contributions of the dissertation were (1) exploring whether teachers, the main gatekeepers in the school system, were biased in their evaluations as a function of students' ascribed characteristics; (2) testing if educational inequalities were concentrated among cognitively weak students, questioning cognitive ability as a valid indicator of merit due to parental strategies of compensation; and (3) providing a causal estimation of the compensatory hypothesis within-families by carrying out comparisons of twins' abilities and tracking outcomes. To carry out the empirical analyses, I used the *Twin Life Study* and the NEPS datasets, applying quasi-causal methods such as twin and school-FE.

This dissertation investigated educational inequalities in Germany. Germany is an interesting case study because it is one of the OECD countries with the highest levels of inequality by SES in academic performance and college enrolment (OECD 2018). Germany is also characterised by low rates of upward educational mobility and social fluidity, and its educational system of early tracking has been indicated as a key element accounting for these high levels of intergenerational inequality cross-nationally. In an international context of technological change leading to the automation of technical jobs and growing income inequalities in the skill premium between the highly- and lowly educated-groups, early tracking into vocational or academic education becomes very consequential for ensuring equal opportunity. This is especially the case in highly-industrialised economies and dual educational systems such as Germany.

Germany also represents an ideal context to test how wealthy kids avoid downward mobility from early in life by assessing the role of skills in educational transitions in a so-called ability-tracking system. Tracking can be considered as an institutional starting gate to evaluate equal opportunity in education (and challenge liberal normative theories) in which teachers are supposed to objectively evaluate students as a function of their ability and behaviour. Therefore, high-SES families may find it particularly difficult to compensate for low ability or negative events for skill formation.

2. Empirical Findings

Results from the first empirical paper (chapter II) on the effect of prenatal health shocks on skill formation showed that lower-BW co-twins have worse academic performance and more behavioural problems than their heavier-BW co-twins. At age 5, there is a causal effect of BW on academic performance and behavioural problems for high- and low-SES families alike. At age 11, the effect of BW on academic skills fades away (or it is reduced) for children of high-SES parents. This pattern of null effects or *compensatory advantage* among high-SES families in the 11-year-old cohort is partially explained by their high absolute level of resources and investments, but not by its relative allocation within-families. The general finding on the detrimental effect of low BW and twins' differences in BW on academic ability is highly consistent with most previous epidemiological literature, while the observed patterns of high-SES families' compensation are consistent with some previous findings from studies in the UK, USA, Chile and Taiwan. Finally, the result on neutral *parental response* to twins' BW differences, or within-family allocation of investments, is in agreement with the limited previous findings on this issue from Chilean twins.

In the second empirical paper (chapter III) studying within-family differences in IQ and tracking outcomes, I found that twins with greater cognitive abilities than their twin siblings enjoy larger transition rates to the academic track of secondary education. This finding aligns with previous research that finds reinforcement patterns for this association in other countries, such as the USA, Mexico, Ethiopia, and Burkina Faso. The positive association between cognitive ability and transition rates to the academic track, generating within-family inequality or reinforcement of abilities, holds for advantaged and disadvantaged families alike. In other words, in contrast to some previous hypotheses and findings, within-family inequality in educational outcomes is not heterogeneous across parental SES. When looking at the heterogeneity of this association across the ability distribution, results show that highly-educated families are not able to compensate for children's low IQ: lower-ability twins at the bottom of the ability distribution show the largest differences in transition rates compared to their relatively more able twin. To my knowledge, this is the first estimation of this association.

The third empirical paper (chapter IV) tested whether there was compensation for low cognitive ability in the transition into the academic track, exploring mechanisms on skill substitution, teachers' bias in assessments and parental aspirations and expectations. I reported four key findings: (1) high-SES students at the same level of cognitive (e.g., competencies in mathematics and German) and non-cognitive skills (e.g., conscientiousness) as low-SES counterparts are considerably more likely to opt for the academic track; (2) in line with the *compensatory hypothesis* and related findings, these socioeconomic inequalities are concentrated among low-performing students (e.g., cognitive competencies in mathematics and German); (3) high-SES students are better able to substitute or compensate for low cognitive skills, having larger returns on non-cognitive skills in the transition to upper secondary; (4) the heterogeneous pattern of skill substitution by parental SES is partially explained by two complementary mechanisms: (i) teachers' bias by students' SES in grading standards and track recommendations; and (ii) higher educational aspirations and expectations of affluent families. Regarding the third finding, as far as I know, there is no previous research on the role of the three-way interaction between parental SES, cognitive and non-cognitive skills in shaping early educational outcomes. With respect to the fourth finding on mechanisms, the results on teachers' bias in grading and track recommendations by students' SES, and on high levels of parental aspirations and expectations net of students' ability, are highly consistent with previous research on inequalities in educational transitions. However, to my knowledge, this paper represents the first application of these mechanisms to explain the compensatory and skill substitution hypotheses.

3. Theoretical Implications and Future Research

What are the theoretical implications of the empirical findings? Generally, I found that (1) high-SES students display more academic ability than low-SES students from age 5, with this SES-gap remaining fairly stable across primary education; and (2) high-SES pupils have more chances of getting ahead on the academic path, over and above ability, than their low-SES counterparts. These general findings align well with skill formation models from economics (first finding), and rational action and cultural reproduction theories on persistent educational inequality from sociology (first and second findings).

The main contribution of the dissertation lied in exploring the heterogeneity of these associations by parental SES to test compensatory theories. Accordingly, I also found that high-SES families are better able to compensate for the negative consequences of prenatal health shocks and low cognitive skills on early skill formation and educational attainment, respectively. These regularities are in line with the predictions of compensatory advantage theories arguing that high-SES families have more resources and incentives, due to risk aversion to downward mobility, to compensate for negative events for status-attainment from early in life.

In particular, in the first empirical paper (chapter II) on BW and skill formation, we tested how the absolute level of resources of high-SES families allows them to deploy high quality and effective investments to compensate for the effects of prenatal health shocks on skill formation. We explored two specific parenting mechanisms that may account for the compensatory advantage hypothesis, such as time spent in cultural activities and emotional warmth. As highlighted by the literature on cultural reproduction theories, developmental psychology, and the findings from interventions targeted at deprived low BW infants in Jamaica and the USA, we found that positive parent-child interactions and cognitive stimulation are important factors for children's skill development. Consequently, we argue that biology is not destiny because (enriched) social environments might offset the detrimental effect of prenatal health shocks on early skill formation.

I tested two additional mechanisms underlying the compensatory advantage hypothesis in the transition into academic secondary education: I tested how cognitively weak but affluent students manage to avoid downward mobility by getting access to the academic track. The analyses testing mechanisms in the third empirical paper (chapter IV) suggest that, on the one hand, teachers perceived high-SES students more favourably by: (i) assigning higher GPA to low-performing high-SES pupils in standardised tests in mathematics and German compared to their low-SES classmates, holding IQ and non-cognitive skills constant; and (ii) by recommending more the academic track to low-performing but striving high-SES students than to low-SES students, controlling for previous parental aspirations and IQ. To infer explicit SES-discrimination from these findings is not clear-cut, but I argue that, if teachers favour characteristics of the student or his/her background that are associated with educational success but not with the student's skills, we can certainly consider this residual effect as a form of teachers' bias towards low-SES families. These results are in line with cultural reproduction theories emphasising the role of teachers as relevant actors in shaping educational inequalities in the school system. However, whether teachers' bias is the result of students' cultural capital, explicit discrimination, or implicit cognitive biases is an unresolved question that future research should address.

On the other hand, high-SES families express considerably higher educational expectations for their low-performing but striving kids than low-SES families. Even when parental educational expectations are supposed to weight in actual barriers and constraints, high-SES parents still express very high expectations for their low ability kids. This finding of *over*-ambitious expectations of advantaged families is consistent with three key mechanisms outlined by rational action theories (RAT) to explain SES-differentials in educational decisions: perceived chances of success, risk aversion to downward mobility, and direct (e.g., transportation costs) and indirect costs.

Five specific findings are somewhat puzzling for the compensatory hypothesis and human capital theories, raising some interesting questions to be addressed in future research. First, the absence of compensatory patterns shown in the second empirical paper (chapter III) on IQ and tracking is in opposition to the compensatory hypothesis, and to human capital theories' revisions hypothesising that high-SES families tend to *compensate* for within-family differences in endowments: siblings with less academic ability achieving the same results as their more endowed siblings. These conflicting results may have something to do with the empirical design adopted. Applying a twin design provides a cleaner causal identification of the association of interest, but external validity is a general concern. When studying within-family differences in twins' school choice, we have to take into account that disadvantaged families are generally more reluctant to opt for the academic track given their low resources, which is especially the case in multiple families with twins, and additional siblings, due to resource dilution. Furthermore, the use of non-verbal IQ as a proxy measure for natural ability is not as directly related to academic performance and GPA as test scores or standardised tests. However, I found that high-SES families still have substantially larger transition rates to the academic track than low-SES families at low, medium or high levels of children's cognitive abilities, even when controlling for parents' IQ. These findings point to the importance of other factors rather than IQ that vary between families and explain educational success, such as non-cognitive skills, test scores, and teachers' bias, as I demonstrated in the third empirical paper (chapter IV).

Second, skill formation models predict that educational interventions or parental investments are more productive in neutralising health shocks or compensating for low levels of skills, during sensitive stages of early child development. Thus, the observed pattern of compensation or null effect of BW among high-SES families at age 11, in comparison to its observed effect at age 5, is not fully in line with theories of human capital formation—even when both birth cohorts are not directly comparable. Otherwise, it could be possible that, as variation in the complexity of learning and skill development arises from early childhood, high-SES parents might have more time to compensate for the detrimental effect of BW throughout pre-school and elementary education.

Third, according to classic microeconomics' theories of intra-household resource allocation, parents allocate resources or investments among siblings depending on preferences, perception of children's endowments, and budget constraints. Generally, reinforcing or compensating parental response to endowments is expected under these models, with varying patterns by family SES. In contrast to previous findings on parental educational investments that reinforce siblings' ability differentials, in the first empirical paper (chapter II), we found no parental response to children's birth endowments whatsoever. This result is in line with previous research using Chilean twins, where a *preference* for equality, neutrality or no parental response is commonly found. Moreover, this "preference" for equality does not seem to vary considerably by SES (see Abufhele, Behrman and Bravo, 2017 for a direct empirical test).

This null result of parental response raises debates about external validity issues when using twin comparisons as a causal research design, and about the actual importance of within-family allocation of resources to explain SES inequalities between families. Using twins as a research design provides advantages in terms of causal identification. However, as acknowledged in chapter II, the twin design also adds theoretical complexity since, even when parents may be willing to differentiate their investments as a function of children's observable health and/or ability, they may find it particularly difficult to do so due to *common goods* and *spillover effects*. Namely, it may be difficult and costly in terms of time and resources to spend independent time with each co-twin or enrol them in different activities (*common goods*), and twins live in the same household and have a very close relationship (*spillover effects*). Besides, many twin families have an extra sibling that is generally older, adding even more complexity to test within-family allocation of resources and investments, an issue that has not been properly addressed in previous research. These issues raise concerns about the trade-off between the pros of causal identification in twin models and the cons of limited external validity that might be addressed by future studies combining sibling and twin samples. Furthermore, as I argue in chapter II, it might be the case that what really matters for explaining inequality dynamics are between-family differences in socioeconomic resources and resource dilution, not within-family allocation, especially so in the case of twins due to the reasons outlined above.

There is an additional limitation of the human capital formation framework in considering all dimensions of parenting as equivalent to investments. This problem distils from classic microeconomic theories of intra-family allocation that formalised parental investments purely in terms of time and money (finite and countable resources) in a rational action cost-benefit setting to maximise children's human capital and earnings. For other dimensions of parenting that are not countable or finite, such as emotional warmth, that recent advances in human capital theories have incorporated following the lead of developmental psychologists

(Attanasio et al. 2020), this framing is admittedly problematic for within-family theories of resource allocation. New theoretical developments integrating human capital formation, intra-household resource allocation by family SES, and different quantitative and qualitative dimensions of parenting are necessary. Overall, the empirical support for the predictions of microeconomics' theories of intra-household resource allocation is very weak at the moment and, up until now, there is little empirical evidence on SES-heterogeneity in preferences, information, and investments within-families. Shedding some light on this mixed evidence of intra-family allocation was one of the main motivations and, I hope, empirical contributions of chapters II and III of my dissertation.

Fourth, the findings of the third empirical paper (chapter IV) on *skill substitution*—low cognitive skills *substituted* by high non-cognitive skills—as a mechanism of compensatory advantage in the transition into secondary education is also puzzling for human capital theories. Generally, skill formation models argue that skills have the property of cross-productivity, so that, for instance, high non-cognitive skills may reinforce cognitive skills and vice versa. Beyond cross-effects, skills can also interact by being complements or substitutes in the production function of learning and educational outcomes. From the concept of skills' cross-productivity, skill formation models implicitly suggest that students with high cognitive skills may complement or reinforce their high learning capacity by combining with high non-cognitive skills: the marginal effect of cognitive or non-cognitive skills is higher among students with previous high ability. However, as shown by empirical findings, the interplay between skills in predicting educational outcomes might depend on parental SES. Thus, whether and how the returns on non-cognitive skills differ across the distribution of cognitive skills, and vice versa, is an open empirical question.

Fifth, in light of the empirical findings, an interesting discussion about the explanatory power and definition of the scope conditions of compensatory advantage theories is granted. One might wonder whether the observed compensatory patterns in Chapter IV are the result of proactive parental investments to compensate for low ability, are just the result of previous parental educational aspirations for their children that are insensitive to performance when expressed as expectations, or are a product of both and additional factors (e.g., teacher's bias). The measurement of academic ability just before the transition into upper secondary education in Chapters III and IV, as in most previous research, might also be misleading because teachers evaluate academic trajectories, not just snapshots, and parents do not discover their kids' ability so late. They can intervene with a plethora of compensatory strategies from birth to age 10. Indeed, testing direct parental response to children's low endowments in terms of health and/or schooling investments might be a better direct test of the compensatory hypothesis, as we did in Chapter II. One could also argue that parental educational aspirations and expectations cannot be really framed as compensatory mechanisms if they do not lead to actual behavioural responses, and they might proxy for other underlying mechanisms (e.g., resources, risk aversion, etc.)

Empirical findings, in line with relative risk-aversion theories, pointed to the key role of parental expectations (and teachers' bias). I still found compensatory patterns (Chapter IV) for low ability—either measured as a snapshot at grade 4 or as a composite from school entry—even when high-SES parents have had 10 years to intervene to boost their kids' academic performance. Thus, the remaining mechanisms to explain compensatory patterns, further than behavioural responses to boost performance (e.g., private tutoring, schooling and health investments), are parental expectations, school search and choice, and teachers' bias in assessments. I tend to think that parental expectations for their children may literally push them to pursue academic education even in the event of low ability. Furthermore, I argue that teachers should be included as main actors in addition to students and their families when trying to explain patterns of compensatory advantage for low ability since they can also influence parental expectations by providing (distorted) signals of their kids' ability (e.g., GPA, recommendations, parent-teacher meetings). These thoughts call for further theoretical developments of the *compensatory advantage model* to test its predictive power in different national contexts and disentangle its underlying mechanisms.

4. Normative and Policy Implications

The dissertation produced novel empirical findings that challenge the liberal conception of equal opportunity and merit in education, with important policy implications. Firstly, children from different social backgrounds do not have the same chances of developing the academic abilities considered as main indicators of merit in the educational system. Inequalities start to gestate in the womb, so that BW, an indicator of child perinatal health, developmental potential, or natural assets in the Rawlsian vernacular, has a long-term differential effect on children's developmental opportunities by socioeconomic circumstances at birth. This finding illustrates how natural assets or endowments interact with social environments in shaping unequal opportunities to develop academic merit from the starting gate of life.

The scarce available evidence shows that the most-effective interventions to mitigate the negative consequences of poor perinatal health—the *Infant Health and Development Program* and the *Jamaican Early Childhood Development Intervention* (McCormick et al. 2006; Walker et al. 2010)—were based on intense psycho-social stimulation by parents and trained professionals (weekly home visits or centre-based) among LBW infants during the first three years of life, when the brain and central nervous system are in dramatic development. The main aim was to improve the quality of mother-child relationships in deprived households by schooling parents on the special needs of LBW kids in terms of cognitive and emotional development, and by providing nutritional supplementation and paediatric follow-up care. Similar policies implemented at the national level seem promising avenues to ameliorate socioeconomic inequalities and foster higher levels of equality of opportunity from birth.

Birth weight, though, might not be a suitable target policy variable if it is proxying for other underlying prenatal conditions and we cannot identify its specific aetiology (Conti et al. 2018)—BW only proxies for the late prenatal environment and most LBW children are pre-term but normal for gestational age. Thus, "while some interventions may indeed succeed in both raising birth weight and improving health outcomes, others may only be effective in raising birth weights, with little or no effects on health (Almond et al. 2005:1074)."

Secondly, the observed SES-inequalities in educational transitions over and above ability represent a waste of academic potential for disadvantaged students, compromising upward social mobility and economic growth in post-industrial societies. The specific findings on compensatory advantage—largest SES inequalities among low ability students—in transition rates to academic secondary education pose a serious challenge to liberal normative theories of equal opportunity that evaluate merit as the sum of natural ability plus effort, with teachers being the gatekeepers or evaluators of merit in the school system. Motivational popular culture clichés and serious empirical findings claiming that *hard work beats talent*—promoting the powerful role of perseverance in compensating for the negative event of low cognitive ability—only seem to be fulfilled for privileged students in the German educational system.

These findings put into question the legitimation of the German system of early-ability tracking as a starting gate based upon selection on meritocratic criteria. I contended that the design of the German tracking system is implicitly inspired by the liberal conception of equal opportunity where a principle of fairness in developmental opportunities is thought to be provided for all children throughout public pre-school and elementary education before the starting gate (e.g., tracking recommendation). The tracking system reflects SES inequalities in academic skills in a race that begins much earlier, even before birth, and strongly selects according to students' ascribed characteristics over and above ability. Thus, the German system of early tracking works instead as a bottleneck that hinders upward mobility through college, where "children of less advantaged origins need to show substantially more 'merit'—however understood—than do children from more advantaged origins in order to enter similarly desirable [educational] positions (Breen and Goldthorpe 2001:82)."

As shown by the Swedish and Finish comprehensive reforms (Meghir and Palme 2005; Pekkarinen et al. 2009; Pekkala Kerr et al. 2013), there is a positive causal effect of erasing early tracking on equality of opportunity, raising intergenerational income mobility over time. Besides, comparative research on educational systems continuously highlights that early-tracking leads to larger socioeconomic inequalities in educational achievement and lower mobility rates in comparison to comprehensive systems (Bol and Werfhorst 2013). Thus, given this evidence in addition to my findings clearly indicating that the German system does not select so stringently on ability as is commonly assumed by academics and the general public, I argue that early tracking should be eliminated if a better allocation of talent and higher rates of equality of opportunity are to be achieved in Germany, as well as in other countries with similar educational systems.

Sociological RAT engage in a normative and policy-implicating debate about the design of educational systems. Rational action theories recommend targeting choice mechanisms to boost equality of opportunity in education, as they argue that cross-national disparities in educational inequality vary as a function of choice, so inferring that SES-gaps in academic ability would be less malleable by social reform. Instead, I argue that social policy to reduce educational inequality should pay special attention to improve developmental opportunities. Bridging SES-gaps in ability may decrease the share of students that were not able to make more ambitious educational choices due to objective or formal insufficient academic ability (e.g., institutional criteria for grade retention), or perceived chances of success (e.g., extreme caution exhibited by the working-classes).

I also found that non-cognitive skills are as important as cognitive skills (e.g., competencies in mathematics and reading) in predicting early educational attainment, and that fluid or non-verbal IQ, the most biologically-driven measure of ability, explains little variation in track choice. These findings are inspiring to mitigate educational inequalities through early interventions boosting skills that are malleable, fundamental for learning, and would not have developed without interventions (Bailey et al. 2017). Disadvantaged students may especially benefit from the stimulation of their potentials, and the postponement of tracking until later ages might be of particular help to allow late bloomers to catch up.

In light of the findings of this dissertation, it is neither possible nor desirable to separate achievement from ascription, nature from nurture, or ability from choice. Thus, liberal normative theories and RAT that assume circumstances and ability, or ability and choice, respectively, to be independent in shaping educational inequalities are seriously challenged. As shown throughout this dissertation, through the complex interplay between biology, social environments, and skills, educational inequalities originate from the starting gate of life in an interactive dynamic of accumulating small (dis)advantages that materialise at critical junctures of the status-attainment process. Thus, paraphrasing George Orwell's 1984 *newspeak* slogan, I shall conclude by claiming that, since *nature is nurture*, *merit and ability are ascription*, and *choice is circumstance*, we should rebuild the project of equal opportunity by "opening up a broader range of opportunities for people to pursue paths that lead to flourishing lives (Fishkin 2014:83)."

References

Abufhele, Alejandra, Jere Behrman, and David Bravo (2017). Parental preferences and allocations of investments in children's learning and health within families. *Social Science and Medicine*, *194*:76–86.

Almond, D., K.Y. Chay, and D.S. Lee (2005). The Costs of Low Birth Weight. *The Quarterly Journal of Economics*, 120(3): 1031-1083.

Attanasio, Orazio, et al. (2020). Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia. *American Economic Review*, 110 (1): 48-85.

Bailey, Drew, Greg J. Duncan, Candice L. Odgers, and Winnie Yu. (2017). Persistence and Fadeout in the Impacts of Child and Adolescent Interventions. *Journal of Research on Educational Effectiveness*, 10(1):7–39.

Bol, T., and Werfhorst, H. G. Van De. (2013). Educational Systems and the Trade-Off between Labor Market Allocation and Equality of Educational Opportunity. *Comparative Education Review*, 57(2): 285–308.

Breen, R., and Goldthorpe, J.H. (2001). Class, Mobility and Merit: the Experience of two British Birth Cohorts. *European Sociological Review*, 17(2): 81–101.

Conti, Gabriella, et al. (2018). "Beyond Birthweight: The Origins of Human Capital." *Institute for Fiscal Studies Working Paper*:18-30.

Fishkin, J. (2014). *Bottlenecks: A New Theory of Equal Opportunity*. Oxford: Oxford University Press.

McCormick, Marie C., et al. (2006). "Early Intervention in Low Birth Weight Premature Infants: Results at 18 Years of Age for the Infant Health and Development Program." *Pediatrics*, *117*(3):771LP –780.

Meghir, C. and Palme, M. (2005). Educational Reform, Ability, and Family Background. *American Economic Review*, 95(1):414-424.

OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/9789264073234-en.

Pekkala Kerr, S., Pekkarinen, T., and Uusitalo, R. (2013). School Tracking and Development of Cognitive Skills. *Journal of Labor Economics*, 31(3): 577–602.

Pekkarinen, T., Uusitalo, R., and Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, 93(7-8): 965–973.

Walker, Susan P., et al. (2010). "The effect of psychosocial stimulation on cognition and behaviour at 6 years in a cohort of term, low-birthweight Jamaican children." *Developmental Medicine Child Neurology*, *52*(7):e148–e154.