



UvA-DARE (Digital Academic Repository)

Not on the same track?

Tracking age and gender inequality in education

Scheeren, L.

Publication date

2020

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Scheeren, L. (2020). *Not on the same track? Tracking age and gender inequality in education*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 1

Synthesis

1.1 Introduction

Even though for many years, men attained higher levels of education than women, the male advantage evaporated and turned into a disadvantage during the last part of the twentieth century in almost all Western countries (Buchmann, DiPrete, and McDaniel 2008; OECD 2011; Vincent-Lancrin 2008). This reversal of the gender gap, also known as the rise of women (DiPrete and Buchmann 2013), is unlikely to disappear any time soon and is suggested to grow even further in the upcoming years (Vincent-Lancrin 2008). As a consequence, a growing body of literature analyses the lagging educational performance of boys (e.g., Epstein 1998; Francis 1999; Mac an Ghaill 1994; Hadjar, Backes, and Gysin 2015; Heyder and Kesels 2017; Martino and Kehler 2006; Morris 2008). The usage of terms such as “the boy problem” or “the boy crisis” by these studies highlights an increasing concern about boys’ educational outcomes in the public discourse.

Although the rise of women took place in almost all Western countries, substantial differences exist. The proportion of women in higher education is, for instance, particularly large in Nordic countries, but substantially lower in some of the Continental countries (Vincent-Lancrin 2008). Not only the magnitude of the gender gap, but also the moment at which women caught up with men varies greatly. The timing was, for instance, much earlier in Nordic areas relative to Anglo-Saxon or Continental areas (Ballarino, Meschi, and Scervini 2013). Even though gender inequality in education is high on the research agenda because of its potential implications for labour markets, marriage markets, family formation, health, social stereotypes etc. (Esteve, García-Román, and Permanyer 2012; Oreopoulos and Salvanes 2011; Van Bavel 2012; Vincent-Lancrin 2008), existing research is not yet able to provide a sufficient explanation for the cross-country and temporal differences in the gender gap in education.

To understand these cross-country and temporal differences, we need to study how gender differences in performance, aspirations, attitudes, behaviour and development are expressed within and affected by different social contexts. A growing body of literature examines the relationship among the organization of society, its institutional features, and gender disparities in education (Becker, Hubbard, and Murphy 2010; Charles and Bradley 2002; DiPrete and Buchmann 2006; Goldin, Katz, and Kuziemko 2006). The main focus of these studies has been on features related to expectations about the life course trajectory and (economic) returns to education, such as the labour market structure and gender-ideological climate. However, of all the institutional features that may be important, I expect to see a particularly strong influence from the institutional context that directly shapes the educational pathways of young people: the educational context.

The relationship between the setup of education systems and gender inequality has rarely been examined, despite educational arrangements being high on the political as well as research agenda because of their crucial relationship with educational attainment and educational disparities by social background and ethnicity (e.g., Hadjar and Gross 2016; Pfeffer 2008; Schütz, Ursprung, and Wössmann 2008; Van de Werfhorst and Mijs 2010). Many countries have adapted the institutional setting of the education system in an attempt to improve educational levels and equalize educational opportunities (Braga, Checchi, and Meschi 2013); potentially, this can also explain cross-country and temporal gender differences in education.

The aim of this dissertation is to study whether a specific setup of the education system works better for girls than for boys or vice versa. Characteristics of the educational institutional context potentially reduce or reinforce gender disparities in education. In this dissertation, I focus on one educational system characteristic: the age at which students are selected into separate educational trajectories, also known as the tracking age. As gender disparities in performance, development and behaviour are not constant across the school career, I expect that the age at which students are tracked plays an important role in educational gender inequality. To what extent is there a specific critical age at which tracking enlarges or reduces gender inequalities in education? By studying various educational outcomes at different moments in the educational career across a large number of countries and birth cohorts, I am able to examine how gender gaps in achievement and attainment develop differently in societies with different tracking ages. Studying how the organization of the education system is related to gender inequality in various contexts will not only inform us about how gender disparities are shaped by the setup of the education system but also help us understand why gender gaps are large in some societies and small in others.

1.2 Gender in context

Especially since the second half of the twentieth century, major steps towards gender equality have been taken, a trend that has been referred to as the gender revolution (England 2010). Women's enrolment in higher education, their attained level of education and their participation in the labour market have increased tremendously (DiPrete and Buchmann 2013; Goldin, Katz, and Kuziemko 2006; OECD 2011; Vincent-Lancrin 2008). Moreover, educational and occupational segregation have decreased (England and Li 2006; Jacobs 1989; Nermo 1996; Weeden 1998). As a consequence, the gender wage gap has also decreased substantially (Charles 2011).

A combination of a wide array of factors has facilitated this gender revolution. For instance, the wide availability of birth control has enabled women to delay child birth (Bailey

2006). Cultural change has resulted in the increasing adoption of gender-egalitarian norms and values (Brewster and Padavic 2000; Inglehart and Norris 2003). The spread of anti-discrimination laws and regulations made gender discrimination in education and the labour market illegal (Burstein 1989). Educational expansion, a growing demand for educated workers and a growing demand for female labour, e.g., due to growth of the service sector, has changed women's qualifications (Akbulut 2011; Becker, Hubbard, and Murphy 2010; OECD 2005). In addition, welfare states have changed, adopting more family-friendly policies that permit or even encourage women to combine family responsibilities with paid work (Chang 2000; Esping-Anderson 1990, 1999; Gornick and Meyers 2003; Mandel and Semyonov 2006).

Despite tremendous progress towards gender equality, these trends have been suggested to be uneven in many ways. First, even though increasing equality has been achieved in some social institutions, e.g., political systems, educational systems and labour markets, inequality is more persistent in other domains, e.g., domestic work (Breen and Cooke 2005). Second, the trends appear to be lopsided, as progression towards gender equality is almost fully due to changes made by women, not men. This has also been referred to as asymmetric change (England 2010). Persistent cultural and economic devaluation of traditionally female-type characteristics and activities, provides men with little incentive to make gender-atypical decisions (England and Li 2006; Levanon, England, and Allison 2009). Women, on the other hand, can gain status and money when moving into male-type activities. As a consequence, remedies for gender inequality today mostly involve prescriptions for women to act more like men, not the other way around.

Third, the trends seem to be uneven because despite women's increased participation in labour markets and educational systems, gender segregation within these institutions remains strong. Compared to vertical gender segregation, horizontal gender segregation, i.e., the disproportionate concentration of men or women in some educational fields or occupational sectors, has proved to be greater and more stable over time (Barone 2011; Bradley 2000; Charles 2005; Charles and Grusky 2004). Persistent gender-specific decisions translate into gender differences in the distribution across fields of study and occupations, with women typically ending up in jobs that are less lucrative in terms of income and status (Gerber and Cheung 2008; Roksa 2005). Increasing gender egalitarianism in the Western world, specifically liberal egalitarian ideals, according to which women are full citizens deserving of rights equal to those of men, has encouraged women to enrol in higher education and attain at least equal levels of education (Charles 2011). As a consequence, vertical gender segregation has decreased. However, these liberal egalitarian ideals easily pair with persistent gender essentialist beliefs, i.e., the belief in innate and fundamental differences between men and women in skills, desires and interests (Charles and Bradley 2002, 2009). These beliefs suggest

that even though men and women are equal, they are different. Therefore, some fields and occupations are believed to be more suited for men than for women and the other way around, which discourages horizontal gender integration. As a consequence, horizontal gender segregation remains strong.

These uneven trends also relate to the fact that what gender means and how gender inequality manifests differ between various social groups, which highlights the importance of examining the intersection of gender with other social categories. The concept of intersectionality, which was first coined by Crenshaw (1989) and is essential to feminist scholarship, refers to the interaction of multiple identities (Davis 2008; McCall 2005). Intersectionality theory argues that social positions are relational, and because of their multidimensionality, gender should not be examined in isolation but in interaction with other social categories, such as social class, race, ethnicity and sexuality. For instance, several studies have shown that there are class differences in the trend towards gender equality in education. Gender segregation in fields of study has been shown to decrease much faster among college graduates and among children from higher social classes, and desegregation has only occurred in middle-class jobs, not in working-class jobs (England 2010; Van de Werfhorst 2017). Both economic and cultural factors account for this context-specificity of gender inequality. Economic explanations suggest that people only transgress gender boundaries when upward mobility within those boundaries is impossible (England 2010). For women from higher social origins, upward mobility will be more difficult within the gender boundaries than for women from lower social origins. Cultural explanations suggest that norms and values related to gender roles are shaped by a person's environment. Conceptions of masculinity and femininity are, thus, context-specific and, for instance, differ between SES groups (Epstein 1998; West and Zimmerman 1987; Willis 1981).

In addition to these uneven trends, studies also emphasize how some trends have stalled. After decades of progress towards gender equality, in many social institutions, degendering has come to a halt. Since the mid-1990s, trends towards more egalitarian attitudes on gender roles have stagnated (Cotter, Hermsen, and Vanneman 2011; Van Egmond et al. 2010). Additionally, with respect to fields of study, labour force participation, occupational segregation, managerial integration and earnings, trends towards gender equality have stalled (Cohen and Huffman 2007; Cohen, Huffman, and Knauer 2009; Cotter, Hermsen, and Vanneman 2004; England and Li 2006; Percheski 2008; O'Neill 2003; Van de Werfhorst 2017). Persistent gender essentialist beliefs impede progress towards gender equality. Studies show that even in the most egalitarian and economically developed countries, a strong emphasis on individual self-expression in combination with economic security and gender essentialism results in persistent gendered decisions (Charles and Bradley 2009). Thus, despite dramatic changes towards equality, some

gender inequalities are much more resilient than others, and we are still far from complete equality between men and women.

1.3 Gender inequality in education

In modern Western society, the social positioning of individuals is strongly based on educational qualifications (Baker 2014). As a consequence, a large body of literature tries to describe and elucidate educational gender inequalities. During the gender revolution, women not only became equal with men but surpassed them in terms of most educational outcomes. Only on test scores in mathematics are boys consistently shown to obtain higher scores than girls, but even this difference is relatively small and in some countries now non-existent (Marks 2008; Penner and Paret 2003; OECD 2011; Stoet and Geary 2013). On almost every other indicator of educational performance, girls outperform boys. Studies have shown that girls receive higher test scores than boys in reading and, for decades, have earned better grades in all major subjects, including mathematics and science (Buchmann, DiPrete, and McDaniel 2008; Duckworth and Seligman 2006; Marks 2008; OECD 2015b; Stoet and Geary 2013). Moreover, even though women have long been disadvantaged in education compared to men, currently women are more likely to enter and obtain a degree in higher education and, in general, attain higher levels of education in almost all Western countries (DiPrete and Buchmann 2013; Goldin, Katz, and Kuziemko 2006; OECD 2011; Vincent-Lancrin 2008). However, as discussed above, horizontal gender segregation remains strong (Barone 2011; Bradley 2000).

Gender differences in the education system are not only about educational outcomes but also about differences in behaviour, development, learning, expectations and aspirations. Prior studies have shown that boys demonstrate more developmental disabilities, more anti-social and disruptive behaviour, lower school engagement, less effort and a less positive orientation to learning (Buchmann, DiPrete, and McDaniel 2008; Downey and Vogt Yuan 2005; Entwisle, Alexander, and Olson 2007; Younger, Warrington, and Williams 1999). Moreover, girls are shown to be ahead of boys with respect to non-cognitive skills, also known as social and behavioural skills, such as goal-directed behaviour, organization, task persistence, self-discipline, ability to cooperate and attentiveness (Anderson et al. 2001; DiPrete and Jennings 2012; Duckworth and Seligman 2006; Jacob 2002; Keulers et al. 2010). Overall, girls consistently display higher levels of pro-school behaviour than boys (Farkas et al. 1990).

All the above-mentioned gender disparities are not just caused by biological differences between men and women. Studies show, for instance, that there are no meaningful differences between men and women in general intelligence and cognitive abilities (Goriounova and

Mansvelder 2019; Halpern 2000). Even if biological factors help explain gender differences in educational outcomes to some extent (Penner 2008), they do not help us understand the substantial cross-national and temporal variability in the gender gaps. This suggests the pivotal role of the environment.

Various country-level features have been proposed to explain this contextual variability. The gender-ideological context is suggested to shape expectations about the roles women can take on later in life and has been shown to affect women's educational aspirations, educational attainment and field-of-study decisions (Buchmann, DiPrete, and McDaniel 2008; Charles and Bradley 2002, 2009; Goldin, Katz, and Kuziemko 2006; McDaniel 2010). Other contextual-level features that are suggested to affect women's educational decision-making have also been shown to affect gender differences in education, such as the labour market structure (Akbulut 2011), total costs of education and (economic) returns to education (Becker, Hubbard, and Murphy 2010; DiPrete and Buchmann 2006; Goldin, Katz, and Kuziemko 2006). The reasoning behind these factors is mostly that boys and girls make decisions on the basis of expectations about their future, i.e., whether it is rational for them to invest in education based on the probability that they will work later in life. Although this rational decision-making process is able to explain part of the cross-country and temporal differences in the gender gap in education, gender disparities cannot fully be explained by country-level variation in the distribution of the costs and benefits of education.

In this dissertation, I point to the important role of the organization of the education system. As the setup of the education system directly affects the educational attainment and educational opportunities of various groups of students in society (Hadjar and Gross 2016; Van de Werfhorst and Mijs 2010), it is puzzling that this contextual-level feature has received so little attention in the literature that tries to explain cross-national and temporal variation in the gender gap.

1.4 The role of the education system

The process of education and schooling is embedded within the educational institutional structure. The setup of the education system plays a central role in levels of schooling as well as educational inequalities along various axes, e.g., social background and ethnicity (Hadjar and Gross 2016; Van de Werfhorst and Mijs 2010). Educational systems shape educational pathways, which can be advantageous for some individuals but limit other people's educational opportunities.

Various features of the educational system have been shown to be important for educational outcomes. Studies have found that the provision of pre-primary education and a

longer duration of compulsory education increase overall levels of education and decrease educational inequality by social background (Braga, Checchi, and Meschi 2013). Smaller educational disparities are also found in educational systems that have a high level of standardization (Van de Werfhorst and Mijs 2010). Furthermore, educational systems characterized by high levels of educational tracking are related to larger inequalities by social and ethnic background (Horn 2009; Lavrijsen and Nicaise 2015; Marks 2005; Van de Werfhorst and Mijs 2010) and lower levels of educational performance and attainment (Hanushek and Wössmann 2006; Brunello and Checchi 2007).

The potential relationship between educational institutional features and gender disparities in education has been mostly overlooked (for exceptions, see Hadjar and Buchmann 2016; Pekkarinen 2008; McDaniel 2010; Van Hek, Buchmann, and Kraaykamp 2019). As the educational institutional context directly affects the educational opportunities of young people while they are still in school, its setup might also be highly consequential for gender disparities in education and might have a significant role in explaining the varying levels of gender inequality across various contexts. In this dissertation, I argue that the organization of educational systems also structures (different) educational pathways for men and women, causing gender divergences in various educational outcomes.

I concentrate on one education system characteristic, tracking age, which is the age at which students are selected into separate educational trajectories on the basis of (cognitive) abilities, interests and aspirations. More specifically, I focus on the age at which between-school tracking occurs. This is the most institutionalized form of tracking and indicates that students of different academic abilities have diverged into completely separate curricular programmes for multiple years (Chmielewski 2014; Maaz et al. 2008). Generally, students are selected into different tracks during secondary education, but the age at which this occurs varies considerably among educational systems (Braga, Checchi, and Meschi 2013). In the Western countries that I investigate in this dissertation, the tracking age ranges from the age of 10 to the age of 16. Germany, Hungary and the Netherlands are examples of countries with a selective system that tracks students at a young age. The United States, Sweden and Spain are examples of countries with a comprehensive system, where students stay in untracked classes much longer (until the age of 16).

The reason I focus on the tracking age is twofold. First, track positions are highly consequential for future educational options. After tracking, students are separated into distinct curricular trajectories that are hierarchically ranked, with the (pre-)vocational tracks being less prestigious than the general/academic tracks. These tracks determine students' future educational options, for instance whether they are allowed to attend university. Especially because track positions have consequences for future educational possibilities, small

inequalities in track placement can translate into large inequalities in later educational outcomes.

Various studies have shown that in early-tracking systems, low-SES students are more likely to transfer to less prestigious, non-academic tracks (Boone and Van Houtte 2013; Buchmann and Park 2009; Caro et al. 2009; Parker et al. 2016). As the track position creates path dependency, inequalities in track positions result in different educational careers among different SES groups and translate into socioeconomic inequalities in future educational as well as occupational outcomes (e.g., Brunello and Checchi 2007; Dustmann 2004; Guyon, Maurin, and McNally 2012; Marks 2005; Van de Werfhorst and Mijs 2010). With respect to gender, we know that boys and girls follow different educational careers and end up with different credentials. Potentially, we can understand these different careers better if we identify the gender differences that are already created when tracking decisions have to be made.

Second, as differences between boys and girls are not constant across educational careers, I expect the timing of important educational decisions to be highly important for gender differences in education. Previous research has shown that gender differences in behaviour, non-cognitive skills and educational performance are not stable across school careers. For instance, women are shown to have an advantage in social and behavioural skills, which grows over time (DiPrete and Jennings 2012). Female-favourable gaps in educational performance are also shown to widen over time (Dekkers, Bosker, and Driessen 2000; DiPrete and Buchmann 2013). Moreover, during secondary education, constructions of masculinity and femininity in relation to schooling become more pronounced; therefore, working hard in school and attaining a masculine identity are becoming increasingly incompatible (Frosh, Phoenix, and Pattman 2002; Swain 2005). The fact that all these differences are not constant over the life course suggests that the timing of important educational decisions is essential for gender differences in education.

The age at which between-school tracking occurs is only one feature of a larger educational institutional characteristic, namely, the level of external differentiation (Bol and Van de Werfhorst 2013). Even though the tracking age is arguably the most important feature that determines the level of external differentiation in an educational system, additional features, such as the number and length of different tracks, are also part of this larger concept. Therefore, various studies measure tracking in a different way. For instance, Bol and Van de Werfhorst (2013) construct an index of tracking based on three indicators that all emphasize a different facet of educational tracking: tracking age, the number of available tracks for 15-year-olds and the percentage of the total curriculum that is tracked. Blossfeld et al. (2016) argue that there is extensive cross-national variation in educational differentiation during secondary education and divide Western countries into four different secondary education models: the early tracking

model (e.g., Germany), the Nordic inclusive model (e.g., Sweden), the individual choice model (e.g., United Kingdom) and the mixed tracking model (e.g., Italy). These studies show that even though some countries, such as Sweden and the United Kingdom, have the same tracking age, their tracking methods are rather different.

In this dissertation, I focus only on the age of first selection. The reason for this choice is threefold. First, theoretically, I expect that it is specifically the age of first selection that is consequential for gender differences in educational outcomes. Gender differences are not stable over time, highlighting the specific importance of the age at which decisions have to be made. Potentially, tracking at a specific critical age enlarges or reduces gender inequalities in education. Second, methodologically, as we focus on a large number of countries and time periods, data on the tracking age are more reliable and easier to compare among different societies than other measures of educational differentiation. Third, empirically, prior research that compares some indicators of external differentiation shows that the tracking age is the key indicator explaining inequality (Horn 2009; Pekkarinen 2008).

As suggested above, the female-favourable gap in educational performance and non-cognitive skills appears to increase during secondary school (Dekkers, Bosker, and Driessen 2000; DiPrete and Buchmann 2013; DiPrete and Jennings 2012; Sammons 1995). Therefore, I expect that educational decisions made later in the school career benefit women and, thus, that a higher tracking age favours women's educational outcomes. To study how the male-female gap in education developed differently in societies with varying tracking ages, in this dissertation, I focus on several educational outcomes at different moments in the educational career across a large number of countries and birth cohorts/time periods.

First, I examine whether the female advantage in education is indeed larger in societies with a higher tracking age. More specifically, I study whether educational reforms in the tracking age affect gender differences in educational attainment (chapters 2 and 3) as well as gender differences in the field of study (chapter 3). Potentially, different mechanisms result in an increased female advantage in early- and late-tracking systems, which indicates that the mechanism at play depends on the institutional setting. To study this theoretical relationship between institutions and mechanisms, I explain variations in the mechanisms through which early and late tracking enhance gender inequality to the advantage of women. To test these mechanisms, I examine, with the use of a longitudinal approach, the extent to which gender gaps in educational achievement and attainment develop differently over the educational career in early- and late-tracking systems (chapter 4). Lastly, I focus on the intersection between gender and social origin and argue that the impact of educational tracking is different for boys and girls from different social backgrounds. I test whether there are stronger advantages or

disadvantages in education related to the tracking age at the intersection of these two social categories (chapter 5).

1.5 Research design

Identifying the impact of system-level variables, such as educational tracking, is rather difficult. To determine the effect of such macro-level features, most studies (have to) rely on cross-country variation for identification (Hadjar and Buchmann 2016; Van Hek, Buchmann, and Kraaykamp 2019). The main issue with this identification strategy is that it strongly relies on the assumption that one can control for all important country-level differences, such as cultural and political traits. Unobserved confounders can strongly bias the results. To overcome this issue, other scholars have used a before-after approach, which focuses on educational reforms within one country (Pekkarinen 2008). However, the generalizability of the results of these studies is problematic, as the effect of a reform is difficult to disentangle from the whole institutional setting of that specific nation. To address these empirical issues, in this thesis, I triangulate between two different identification approaches to examine the effects of educational tracking. One focuses on educational reforms in various countries, and the other adopts a differences-in-differences approach.

Identifying the effect of tracking reforms

In the first two empirical chapters, I study the impact of educational tracking on gender differences in education with the use of educational reforms. I focus on differences in tracking age across countries as well as differences over time, i.e., before and after reforms, exploiting a within-country over-time design. More specifically, I examine whether policy reforms concerning the tracking age affect gender differences in educational outcomes with the use of a cohort design, which means that I compare the educational outcomes of birth cohorts that experienced different tracking ages during the time they were in education in various countries. A great advantage of this design is that it allows me to control for all unobserved time-constant country-specific features as well as unobserved general time trends. Moreover, as many countries and time periods are studied, these results are easier to draw generalizations from.

To implement such a model and examine whether educational reforms affect gender differences in various countries, dynamic contextual-level data on educational tracking are necessary. The educational reform dataset constructed by Braga et al. (2013) provides me with information on educational reforms concerning the tracking age in various European countries between 1929 and 2000. This information, namely, the age at which educational systems split students into separate trajectories within all these countries for the whole time period, enables

me to create a continuous measure of the tracking age, which ranges from the age of 10 in the earliest-tracking countries to the age of 16 in the latest-tracking countries.

In chapter 2, I examine whether educational reforms affect gender differences in educational attainment. In this chapter, I combine contextual-level information on the tracking age with individual-level information on final educational attainment, which I derive from multiple rounds of the *European Social Survey* (ESS). The reason I choose to focus on educational attainment is fourfold. First, this measure is part of several large cross-national datasets, which provide me with information on the educational attainment of many individuals from various birth cohorts and countries. Most previous research on the impact of features of the education system focuses on educational achievement (test scores) in specific subjects (e.g., Bedard and Cho 2010; Hanushek and Wößmann 2006; Horn 2009; Lavrijsen and Nicaise 2015; Van Hek, Buchmann, and Kraaykamp 2019); however, data on achievement are only available for recent cohorts. As many reforms in the tracking age happened decades ago, I prefer to use a wider timeframe.

Second, I expect that the impact of educational tracking might be greater for educational attainment in particular than for educational competence. Selection into educational tracks is based not only on students' cognitive abilities but also, among other things, on students' ability to show their educational potential, their grades, and their aspirations. This indicates that track placement selects students with not only the best cognitive skills but also the best social and behavioural skills to advance to higher levels of education. Therefore, the impact might be particularly large for educational outcomes that are, next to cognitive skills, also highly related to social and behavioural skills, such as educational attainment.

Third, focusing on final educational attainment allows me to examine potential long-term consequences of educational tracking. Small short-term consequences might translate into large long-term consequences, highlighting the importance of also examining longer horizons. Lastly, as final educational attainment has important implications for, among other things, wages and labour markets, examining gender inequality in educational attainment is important for our understanding of gender inequality in the social positioning of individuals in society.

In chapter 3, I examine whether educational reforms affect gender differences by field of study and, as a replication of chapter 2, educational attainment. In this chapter, I combine contextual-level information on the tracking age with individual-level information on the field of study and final educational attainment, which I gather from multiple rounds of the *European Union Labour Force Survey* (EU LFS). To understand the impact of educational tracking on gender inequality, it is important that we also focus on horizontal forms of segregation. Gender inequality by field of study has been shown to be greater and less resilient than gender inequality in the level of education (Barone 2011; Charles and Bradley 2002) as well as an

important driver of occupational gender segregation and gender wage gaps (Charles 2010). Hence, to comprehend the potential consequences of educational tracking for gender inequality in labour markets and wages, in this chapter, I study how the tracking age impacts gender inequality by field of study.

One of the main advantages of this identification strategy is that it automatically controls for all important time-constant country-level differences. Nevertheless, time-varying country-level differences can still bias the results. Therefore, in these chapters, I control for several potential time-varying confounders, e.g., gender-ideological climate and labour market demand. Data on these indicators are gathered from different sources, such as the World Databank and the EU KLEMS database. I refer the reader to the empirical chapters for their full descriptions.

Identifying the tracking effect

In the last two empirical chapters of this thesis, I study the impact of educational tracking on gender differences in education with the use of a differences-in-differences (DiD) design (Hanushek and Wössmann 2006; Ruhose and Schwerdt 2016). Within such a design, the focus is not on the specific level of educational inequality but on the changes between two time points in the level of educational inequality. I compare these changes among various countries with different tracking ages. Thus, this strategy identifies the tracking effect by exploiting cross-country variation in the tracking age as well as variation in the evolution of gender differences between two time points. More specifically, in these chapters, I examine changes in educational gender gaps between two time points, for instance, between primary and secondary school, and test whether there are significant differences in these changes in countries that implemented tracking between these two time points and countries that did not. If the evolution of the gender gap between primary and secondary school is different among countries that started tracking between these two time points and countries that did not, this points to tracking affecting gender inequality in education. As this model takes into account the gender gaps that already existed before tracking took place in any country, unobserved (time-constant) heterogeneity between countries is controlled for, and therefore, this model provides me with a more precise estimate of the impact of educational tracking.

In the abovementioned cohort design, I compare birth cohorts that experienced different tracking systems. In this life course design, I examine changes within birth cohorts over the life course. What changes occur after these birth cohorts experience tracking? To implement this model, I need comparable information on educational outcomes at multiple time points, i.e., data before tracking takes place in any country (before the age of 10), data after tracking takes place in early-tracking countries but not in late-tracking countries (around the age of 15),

and data after tracking takes place in all countries (after the age of 16). In line with prior studies using such a model, the tracking age is a dummy variable measuring whether a country has an early- or a late-tracking system, where tracking is considered early when students are streamed into separate trajectories before the age of 15 (Hanushek and Wössmann 2006; Ruhose and Schwerdt 2016).

To examine gender differences at these different time points in early- and late-tracking countries, I combine a large number of high-quality repeated cross-national surveys that assess or survey respondents in several countries at different ages and educational stages. More specifically, I merge several waves of international comparative student assessment data from primary school (PIRLS [Progress in International Reading Literacy Study] and TIMSS grade 4 [Trends in International Mathematics and Science Study]), secondary school (TIMSS grade 8 and PISA [Programme for International Student Assessment]) and early adulthood (PIAAC [Program for the International Assessment of Adult Competencies]). For information on educational attainment, I also merge the ESS in this large dataset.

As comparable information on educational achievement (test scores) in mathematics and reading is available at different (young) ages, these are the main educational outcomes considered in the last two empirical chapters. The advantage of focusing on educational achievement is that the repeated data allow me to investigate how gender gaps develop over the school career and how this differs between early- and late-tracking countries. This provides me with more insight into when and how tracking affects the educational gender gap. Moreover, with these measures, I can examine the immediate consequences of the tracking age for cognitive skills. If such consequences exist, educational tracking not only sorts students selectively into different tracks but also affects students' learning within these tracks. Finally, examining the impact of tracking on educational achievement provides me with information on the underlying mechanisms through which tracking affects gender differences in educational attainment.

A note on gender

If educational systems (or other macro-factors) matter for the size of the gender gap, this illustrates that inequalities between boys and girls are socially, culturally and institutionally shaped. In regard to the social, cultural and institutional explanations for sex differences in careers, it appears natural to use the term gender. The term sex refers to biological features that classify individuals as 'male' or 'female'. The term gender refers to a person's features that are recognized or culturally defined as masculine or feminine by a social world. While this thesis focuses on differences between men and women in a binary way, we use the term gender throughout the entire dissertation. Quite some debate exists about the way to measure gender

and its non-binary, multidimensional character (Bittner and Goodyear-Grant 2017; Smiler and Epstein 2010). A strong overlap in traits exists between boys and girls, and within-group heterogeneity is often larger than the differences between boys and girls. Unfortunately, operationalizing gender in any way other than the binary way is not possible with the datasets that I use in this dissertation. Especially since I focus on a large number of countries and time periods, no dataset that would make possible an alternative operationalization exists. Hopefully, future research will be able to use a finer-grained measure of gender to examine gender differences in education in more detail.

1.6 Four empirical studies

Chapter 2. The gender revolution in context: How later tracking in education benefits girls

In the first empirical chapter of this dissertation, I focus on the role of the educational institutional context in the increasing female advantage in education. More specifically, I examine whether the age at which students are selected into separate educational trajectories, i.e., the tracking age, is related to the male-female gap in educational attainment. Girls have an increasing advantage over boys in non-cognitive skills and educational performance during secondary education, which suggests that educational decisions made later in the school career benefit girls. Therefore, I hypothesize that a higher age of first selection will favour women's educational attainment.

I combine individual-level data on completed years of education from the ESS with unique longitudinal contextual-level data on educational reforms between 1929 and 2000 in 21 European countries and employ a two-level multilevel model including country and cohort fixed effects. The results suggest that educational reforms that postponed the timing of tracking have particularly benefited women's years of education relative to men's, even when various other influential factors are controlled for, such as the gender-ideological climate or labour market demand. This finding shows that the rise of women is not homogeneous across institutional contexts, as some educational systems contribute to women's increasing educational outcomes more than others.

Chapter 3. Female advantage and disadvantage: The role of educational tracking in vertical and horizontal gender segregation in education

In chapter 3, I go beyond vertical gender segregation and examine whether the impact of the tracking age extends to horizontal gender segregation in education. I hypothesize that a higher

age of first selection decreases gender segregation across fields of study. Moreover, in this paper, I replicate the analysis of chapter 2 on vertical gender segregation and, in line with chapter 2, hypothesize that a higher age of first selection is beneficial for women's educational attainment.

Microdata on educational outcomes from the EU LFS are combined with longitudinal contextual-level data on educational reforms concerning the tracking age that were implemented between 1966 and 2006 in 20 European countries. I exploit a within-country over-time design, employing linear and multinomial models with country and cohort fixed effects. My results indicate that policy reforms implementing later tracking are associated with both horizontal and vertical gender segregation in education. I find that later tracking reduces horizontal gender segregation among the higher-educated people, as both men and women become more likely to complete a degree in gender-atypical fields when the tracking age is postponed. Moreover, similar to the results in chapter 2, I find that the postponement of tracking in an educational system contributes to women's advantage in completed years of education, increasing vertical gender segregation.

Chapter 4. Gender inequality in educational performance over the school career: The role of tracking

In chapter 4, I adopt a longitudinal perspective to examine the extent to which educational gender gaps develop differently over the educational career across educational systems that vary in their age of first selection and at what point in the school career tracking is (most) consequential for gender inequality in education. I investigate two different hypotheses. The cumulative advantage hypothesis states that girls' higher likelihood of selecting into higher tracks, as a result of a female advantage in educational performance and non-cognitive skills in primary school, leads to a potential accumulation of advantage for girls in early-tracking systems. The critical age hypothesis states that girls outperform boys to a larger extent later in the educational career, suggesting that later tracking leads to much larger educational gender gaps to the advantage of women.

I test these hypotheses by investigating gender differences in educational achievement and attainment in primary, secondary, and post-secondary education. Individual-level data on educational outcomes, retrieved from PIRLS, TIMSS, PISA and the ESS, are matched with country-level information on the age of first selection. Using a differences-in-differences (DiD) approach, I examine whether gender gaps in educational achievement and attainment evolved differently over the educational career in early- and late-tracking systems while controlling for all cross-national variation in the gender gap that already existed before tracking took place.

The results provide support for both hypotheses, as I find that educational tracking, regardless of the age at which it takes place, is, to some extent, beneficial for girls. This suggests that the story is slightly more complicated, as the mechanism at play depends on the institutional setting. However, the impact of educational tracking is much larger when educational systems track late than when they track early, indicating that tracking at an older age is more consequential for the gender gap in education than tracking at a younger age.

Chapter 5: The differential impact of educational tracking on SES gaps in reading achievement for boys and girls

A large body of literature has shown that later tracking is related to smaller inequalities by socioeconomic origin. In chapter 5, I focus on the intersection between socioeconomic status (SES) and gender and examine whether later tracking reduces social inequality in reading performance among boys and girls. I argue that due to cultural differences between SES groups, the extent to which students engage in pro-school behaviour differs between boys and girls of different social backgrounds. Whereas working-class versions of masculinity emphasize the incompatibility between masculinity and pro-school behaviour, middle-class versions emphasize academic competition and ambition. As a consequence, I expect differences between low- and high-SES boys in pro-school behaviour to increase during high school. Staying longer in untracked classes is, therefore, not likely to decrease SES gaps among boys. As pro-school behaviour is stereotyped as feminine, differences among girls of different SES groups are much smaller. This pro-school behaviour might help low-SES girls compensate for their disadvantaged background if they are provided with enough schooling time in comprehensive classes before they are tracked. Therefore, we hypothesize that staying longer in untracked classes reduces inequality by social origin for girls but not for boys.

Microdata on reading competence during primary school (PIRLS) and secondary school (PISA) are combined with contextual-level data on the tracking age. I employ differences-in-differences (DiD) models to examine whether the impact of tracking on SES-based inequality in reading achievement differs between boys and girls, while controlling for cross-national variation in SES and gender inequality that already existed before tracking occurred. In line with my hypothesis, I find that later tracking reduces inequalities by social origin in reading achievement for girls. Staying longer in untracked classes does not, however, affect SES-based inequalities for boys.

An overview of the four empirical chapters, including the key question, the individual-level data, the employed method and the main results of each, is provided in Table 1.1.

Table 1.1 Overview of the empirical chapters

	Key Question	Individual-level data	Method	Main results
Chapter 2	To what extent does the tracking age affect the gender gap in educational attainment?	<ul style="list-style-type: none"> • ESS 2002, 2004, 2006, 2008, 2010, 2012, 2014 	Two-level multilevel model, including country and cohort fixed effects	Late tracking improves women's completed years of education relative to men's.
Chapter 3	To what extent does the tracking age affect vertical and horizontal gender segregation in education?	<ul style="list-style-type: none"> • EU LFS 2005, 2010, 2015 	Two-level multilevel model and multinomial model, both including country and cohort fixed effects	Late tracking increases vertical segregation to the benefit of women and reduces horizontal segregation among higher-educated people, as both men and women become more likely to complete a degree in gender-atypical fields.
Chapter 4	To what extent do educational gender gaps evolve differently over the educational career in early- and late-tracking systems, and at what point in the school career is tracking most consequential for gender differences?	<ul style="list-style-type: none"> • PIRLS 2001, 2006, 2011 • TIMSS grade 4 1995, 2003, 2007, 2011, 2015 • TIMSS grade 8 1995, 1999, 2003, 2007, 2011, 2015 • PISA 2000, 2003, 2006, 2009, 2012, 2015 • PIAAC round 1 (2008-2013), round 2 (2012-2016) • ESS 2002, 2004, 2006, 2008, 2010, 2012, 2014 	Differences-in-differences (DiD) model	Educational tracking is, irrespective of the age at which it occurs, beneficial for women. However, the positive effect for women is much stronger when tracking occurs later in the school career.
Chapter 5	To what extent is the impact of the tracking age on social inequality in reading performance similar for boys and girls?	<ul style="list-style-type: none"> • PIRLS 2001, 2006, 2011 • PISA 2000, 2003, 2006, 2009, 2012, 2015 	Differences-in-differences (DiD) model	Whereas later tracking reduces SES-based inequality in reading achievement among girls, it does not affect SES-based inequality among boys.

1.7 Five lessons learned

Due to its key role in educational levels and educational inequalities, the setup of the educational system is often at the centre of political debates. Various countries have reformed their educational systems to reduce inequalities related to social and ethnic background or

enhance overall levels of education (Braga, Checchi and Meschi 2013; Hadjar and Gross 2016). In this dissertation, I show that discussions on educational systems cannot ignore gender inequality. Overall, this thesis provides us with five important take-away messages.

First, this dissertation highlights that educational tracking is important in generating educational gender inequalities. In general, I find that later tracking enhances the female-favourable gap in educational achievement and attainment and reduces gender segregation across fields of study among the higher-educated people, as both men and women are more likely to obtain a degree in gender-atypical fields in late-tracking systems. Even when taking into account several other contextual-level features that potentially affect gender differences in educational outcomes, such as the gender-ideological climate and labour market demand, I demonstrate that the age of first selection affects the educational gender gap.

Second, this thesis emphasizes the importance of taking into account gender when focusing on the implications of educational systems, as I provide evidence that the relationship between tracking and socioeconomic inequality is driven by females. Various studies have shown that later tracking reduces educational inequality related to social origin (e.g., Horn 2009; Marks 2005; Van de Werfhorst and Mijs 2010). I find that later tracking reduces SES-based inequality in education among girls but does not affect inequality among boys. This suggests that later tracking enlarges gender inequality because low-SES girls are able to compensate (a bit) for their disadvantaged background in later tracking systems, whereas low-SES boys are not. It is especially low-SES boys who do not benefit from later tracking. This indicates that educational reforms that are suggested to reduce educational inequality only help low-SES girls. As a consequence, low-SES boys fall even further behind after such a reform.

Third, this dissertation shows that educational institutional structures affect educational levels and educational inequalities along various axes, leading to multiple trade-offs. Often, political debates highlight one particular positive or negative consequence of a specific feature of the educational system, but trade-offs exist (Bol and Van de Werfhorst 2013). A first important trade-off that this thesis detects is that even though a higher tracking age diminishes educational inequality in relation to social origin (Horn 2009; Marks 2005; Van de Werfhorst and Mijs 2010), it increases educational inequality by gender; i.e., it enlarges the female advantage in educational attainment. A second trade-off is that, compared to early tracking systems, later tracking improves educational outcomes for both men and women, indicating higher overall levels of education, but the advantage for women is much stronger, resulting in increased gender inequality. A third trade-off is that while later tracking reduces horizontal gender segregation among higher-educated people, it enhances vertical gender segregation. I do not suggest that all different forms of inequality or all the abovementioned trade-offs are equally important. Nevertheless, it is important to be aware that the educational institutional

structure plays a key role in shaping educational attainments and educational opportunities of various groups of students in societies. Reforms of educational systems likely have different implications for all of these different groups.

Fourth, this thesis shows that tracking decisions, irrespective of timing, benefit women. Even though the impact is larger when tracking occurs later than when it occurs earlier, I find that every timing of tracking benefits women's educational outcomes more than men's. This indicates that the story is slightly more complex. The mechanism at play depends on the institutional setting. In early-tracking systems, girls' advantage in educational performance and social and behavioural skills in primary school increases their likelihood of attending higher tracks. Subsequently, their advantage in track placement leads to an accumulation of advantage during the rest of their educational career. In late-tracking systems, the beneficial effect of the tracking moment for girls is much larger, as girls outperform boys to a larger extent later in the educational career. This suggests that late-tracking systems track at a critical age. Even though the second mechanism is stronger, resulting in a larger female advantage in late-tracking systems than in early-tracking systems, both mechanisms are at play. This suggests that boys are disadvantaged by tracking, irrespective of the age at which it occurs. A growing body of literature tries to understand the "boy problem" (e.g., Epstein 1998; Francis 1999; Mac an Ghail 1994; Hadjar, Backes, and Gysin 2015; Heyder and Kesels 2017; Martino and Kehler 2006; Morris 2008) and argues that the disproportionate number of female teachers in primary school, peer effects and gender stereotypes at school might explain boys' lagging educational outcomes relative to girls'. This dissertation highlights that the educational decision-making process and the timing of educational decisions are also part of the explanation.

A final important message of this dissertation is that gender differences across the school career are not constant; therefore, the timing of important educational decisions plays a crucial role in generating educational inequalities. In this dissertation, I focus on the impact of tracking age. However, the results potentially have implications for any intended reform within the educational system related to the selection and differentiation of specific age groups. Hence, the timing of any important transition within the educational system needs to be a contributing factor in the decision-making process.

How to set up the educational system to enhance overall levels of education and diminish inequalities is an important question that many policymakers would like to see answered. The lessons learned from this dissertation provide us with some new insights into the answers to this question. This thesis shows that gender inequality in educational attainment is smaller in countries that track early. However, I also show that early tracking mostly harms women, as it specifically holds back women's educational outcomes, preventing them from reaching their full educational potential in such a system. Furthermore, it specifically harms women from a

low-SES background, as low-SES women are unable to compensate for their disadvantaged background in early-tracking systems. This suggests that, from an equal opportunities point of view, tracking at an early age might not be desirable. Moreover, I find that later tracking improves, at least to some extent, the educational outcomes of men and women. If enhancing overall levels of education is the main goal, tracking early might not be desirable. Nevertheless, the fact that the achievement and attainment of boys relative to that of girls is lower in these late-tracking systems is a point of concern that cannot be ignored and should receive more attention in future research.

1.8 Future research

Given that educational qualifications are essential for the social positioning of individuals in today's society (Baker 2014), it is imperative to understand how specific inequalities are generated and perpetuated. This dissertation highlights the importance of the setup of the education system for gender inequality and, more specifically, shows that the age at which educational systems track affects gender differences in various educational outcomes. The results do, however, also raise new questions. I provide four avenues for future research.

A first important direction for future research, to improve our understanding of the underlying mechanism, is examining gender differences over the school career with the use of longitudinal data. This dissertation provides theory-driven research between countries and suggests that the mechanism at play depends on the institutional setting. I argue that boys and girls differ in their development, show distinct behaviours in class and possess cognitive and non-cognitive skills to varying degrees at different moments in the school career. Therefore, boys and girls react differently and profit differently from specific educational structures. To zoom in and examine gender differences over the whole school career, we are in need of longitudinal data focusing on boys' and girls' skills, behaviour and norms during the educational career. Obtaining more insight into these gender differences and how they change over the educational career will not only help us understand how the tracking age is related to gender gaps in education and why boys are not able to benefit (as much as girls) from late tracking but also how various other contextual-level features, at the family, school or country level, are related to gender differences.

Second, as suggested above, hopefully future research is able to use a finer-grained measure of gender to implement a more direct test of tracking on gender inequality in education. Especially since I argue that constructions of masculinity and femininity in relation to schooling and how this changes over the school career are part of the mechanism underlying the relationship between tracking and gender inequality, measuring gender in more detail can

provide us with great insight. To study more detailed forms of gender inequality within different institutional contexts, we are in need of cross-country data that allow researchers to measure gender as a multidimensional concept.

A third important direction for future research is studying the impact of educational tracking within its wider context. Educational institutional structures do not exist in isolation but are part of a larger education system and exist within a broader institutional context. The impact of educational tracking might be dependent on this context. For instance, the impact of later tracking on gender inequality might be smaller in countries that also have high levels of standardization, as is shown to be the case with SES-based inequality (Bol et al. 2014). Alternatively, the impact of tracking-age reforms on gender inequality might be larger in more gender-egalitarian contexts. Future research should examine possible interactions between educational tracking and other contextual-level features and investigate whether specific combinations are able to diminish several forms of educational inequality and potentially solve some of the abovementioned trade-offs.

A fourth important direction for future research is examining whether and how higher levels of vertical segregation and lower levels of horizontal segregation in education in late-tracking systems translate into gender differences in the labour market. Even though women are becoming increasingly more educated than men, this does not necessarily translate into equality in other areas as well, e.g., income and occupational status (Charles 2000). Moreover, less horizontal segregation in education does not automatically translate into less occupational gender segregation (Smyth and Steinmetz 2008). To understand the consequences of women's higher educational outcomes in late-tracking systems, future research should focus on the relationship between the tracking age and gender inequality in other social domains.