Not on the same track?

Tracking age and gender inequality in education

Scheeren, L.

Publication date
2020

Document Version
Other version

License
Other

Chapter 2

The Gender Revolution in Context: How Later Tracking in Education Benefits Girls[1]

Abstract It is well established that women have surpassed men in educational attainment. The potential effect of educational institutional structures on the "rise of women" has largely been overlooked. In particular, the age of first selection into distinct educational tracks seems important, as the female-favourable gap in educational performance has been shown to widen during the high school years. Using microdata from the European Social Survey and data on tracking age reforms for 21 European countries from 1929 to 2000, we find that reforms promoting later tracking have particularly benefited girls' years of education, net of other influential factors (gender-ideological climate and demand on the labour market). This finding shows that the rise of women is not homogeneous across institutional contexts: some educational systems are more supportive of the enhancement of women's education than others.

[1] A slightly different version of this chapter has been published as Scheeren, Lotte, Herman G. Van de Werfhorst, and Thijs Bol. 2018. "The Gender Revolution in Context: How Later Tracking in Education Benefits Girls." *Social Forces* 97(1):193-220.

2.1 Introduction

Women have long been disadvantaged in education compared to men, while currently they outperform men in educational attainment in many industrialized countries. This "rise of women" (DiPrete and Buchmann 2013) is likely to persist and even grow in the upcoming years (Vincent-Lancrin 2008). Although the reversal of the gender gap in educational attainment has occurred in almost all Western countries (OECD 2011), substantial differences exist. Both the magnitude of the gender gap in education and the time at which women caught up with men vary greatly. Most studies have examined individual-level explanations (e.g., Buchmann and DiPrete 2006), but recent work has studied institutional explanations for the cross-country and over-time differences in women's relative educational attainment. This literature has examined mostly contextual factors that influence expected (economic) returns to higher education, such as changes in gender ideological context or the labour market structure (Becker, Hubbard, and Murphy 2010; Goldin, Katz, and Kuziemko 2006). However, the role of the *educational* institutional context has largely been overlooked. It is unclear whether a specific setup of the educational system works better for girls than for boys or vice versa.

In this article, we study whether and how the age of first selection in the educational system affects the gender gap in educational attainment. In many European educational systems, children are tracked into separate school trajectories much more rigorously than in the United States. Especially in early tracking countries, such as Germany, Austria and the Netherlands, children enter separate schools for multiple years beginning when they are 10 to 12 years old and have different formal opportunities for their later school careers. Other European systems offer comprehensive education until the age of 16, including Scandinavian countries. Interestingly, these system differences may affect boys and girls precisely in the period when the male-female gap in academic performance widens to the benefit of females, namely, the teenage years (DiPrete and Buchmann 2013; Sammons 1995). If girls have increased advantages during the middle and high school years in terms of cognitive and non-cognitive skills that are conducive to further performance, girls may be at a particular advantage if selection happens later in their school careers.

Whereas some evidence indicates that a later age of selection benefits women's educational attainment relative to men's (Jürges and Schneider 2011; Pekkarinen 2008), no study to date has empirically tested this hypothesis for more than one country or over a long period of time. Including several countries and a long time period allows us to examine the relationship more rigorously. Moreover, a large set of countries enables us to exploit multilevel modelling approaches by which we can examine tracking effects both within countries over

time and between countries. In some of these models, we include country and cohort fixed effects to control for all time-invariant country-specific factors and unobserved general time trends that affect all countries.

We combine longitudinal contextual data on educational reforms (Braga, Checchi, and Meschi 2013) from 1929 until 2000 with individual data on completed years of education in 21 European countries, retrieved from the European Social Surveys (ESS). In addition to adding country and cohort fixed effects, we also control for a large number of dynamic contextual variables that test rival explanations for changing gender gaps in education, including gender-ideological climate and demand on the labour market.

To foreshadow our results, our models indicate that later tracking favours women's completed years of education. Even when we take into account the possibility that several factors might have changed in certain countries over time, we find that policy reforms concerning the tracking age affect the male-female gap in attained years of education. This result is not only relevant for educational systems with between-school tracking or countries that have experienced educational reform. What we find is that important educational decisions – about what level or educational programme to follow – have a differential impact on students from different sexes. More specifically, the timing of these decisions has consequences for gender differences in educational outcomes. This is a finding that potentially has implications for any intended reform targeting the educational selection and differentiation of specific age groups.

2.2 Theory

Tracking age and gender inequality

Curricular tracking, also known as ability grouping or educational differentiation, is the extent to which students are allocated into different educational, school classes, or schools on the basis of performance and/or aspirations (Vanfossen, Jones, and Spade 1987). Our study focuses on what has been called "between-school tracking" (Chmielewski 2014), which implies that students are separated for several years and enrolled in different curricular programmes. Moreover, between these demarcated educational trajectories, often labelled the (pre-)vocational and general/academic tracks, it becomes evident which is the more prestigious and which is the less prestigious track (Bol and Van de Werfhorst 2013).

Students are generally streamed into different tracks during secondary education; however, the age at which this occurs differs considerably between educational systems (Braga, Checchi, and Meschi 2013). In this study, we focus on this crucial aspect of tracking: the age at which national educational systems separate students into different programmes (Horn 2009). In more

selective systems, students are tracked into streams during lower secondary education, sometimes as early as age 10. In comprehensive systems, on the other hand, students remain much longer in untracked classes, up to age 16.

While the tracking age is probably one of the most extensively investigated educational system characteristics, it has been rarely studied from a gender perspective. A large body of literature indicates that earlier tracking is associated with increased socioeconomic and ethnic inequality (Horn 2009; Van de Werfhorst and Mijs 2010). Prior research has also demonstrated that a selective system decreases school performance and final educational attainment (Hanushek and Wössmann 2006; Brunello and Checchi 2007). As a consequence, educational systems with early tracking have received a lot of criticism (OECD 2007).

While the focus has thus been mostly on socioeconomic and ethnic inequalities, a few studies have examined the effect of the tracking age on gender differences in education, showing that educational systems with a later tracking age tend to favour girls. Jürges and Schneider (2011), who investigated the German system, found that the female-favourable gap in track recommendations and enrolment was slightly larger when the tracking age was 12 instead of 10. Pekkarinen (2008), who examined a Finnish school reform, showed that when the tracking age was postponed from 10-11 to 15-16, the gender differences in the likelihood of picking an academic track and enrolling in tertiary education became larger. Finally, Hadjar and Buchmann (2016) showed that educational attainment of women was relatively higher in late-tracking than in early-tracking educational systems.

Why would we expect that later tracking has (more) positive effects on the educational attainment of females versus that of males? Prior research has demonstrated that the female-favourable gap in educational performance widens over the educational life course (Dekkers, Bosker, and Driessen 2000; DiPrete and Buchmann 2013). Gender differences in prior educational achievement did not appear to be an important explanation for increasing gender differences in educational achievement during the high school years (Sammons 1995; Machin and McNally 2005), indicating that the gap increases as children grow older. In general, track assignment is based on previous educational performance, as well as expectations about a student's educational potential. In this study, we provide two arguments why the female-favourable gap, in both educational performance and expectations about educational potential, is larger at the end of lower-secondary education than earlier in the educational career.

First, there is a sex difference in the timing and speed of maturation. Generally, during adolescence, girls are developmentally more advanced than boys: girls enter puberty at an earlier age and have more advanced psychosocial and brain development (De Bellis et al. 2001; Giedd et al. 2006). As a consequence, boys temporarily lag behind girls in important non-cognitive skills (Keulers et al. 2010), also referred to as executive functions or social and

behavioural skills. These non-cognitive skills include cognitive self-regulation, including planning and sustaining attention, and emotional self-regulation, such as the ability to control emotional responses and sociability (DiPrete and Jennings 2012). Boys' lower non-cognitive skills during adolescence are reflected, for instance, in higher anti-social and disruptive behaviour in school and lower attentiveness and school engagement in general (Downey and Vogt Yuan 2005; Jacob 2002). Previous studies showed that girls' advantage in non-cognitive skills, which is already present in kindergarten, increases as they progress through school, resulting in large gender gaps in non-cognitive skills during secondary education (Anderson et al. 2001; DiPrete and Jennings 2012; Keulers et al. 2010). As these skills are important for academic choices and performance (Best, Miller, and Naglieri 2011; OECD 2015a), a rising non-cognitive skills gap increases gender gaps in educational performance during secondary education. Moreover, non-cognitive skills are important to signal educational potential (Downey and Vogt Yuan 2005). This suggests that when students have to make educational choices in late-tracking systems compared to early-tracking systems, girls will have a lead over boys in both educational performance and non-cognitive skills. Educational choices made at an older age during secondary education will then likely benefit girls' educational outcomes relative to boys'.

Second, gender identities can also explain gender differences in both social and behavioural skills and educational performance. Boys and girls engage in specific behaviour to accomplish a feminine or masculine identity (Morris 2008; West and Zimmerman 1987). "Hegemonic masculinity", the traditional notion of male identity, means being dominant and even sometimes behaving in a deviant way (Connell and Messerschmidt 2005). Studies have argued that this construction of masculinity promotes boys having "laddish" attitudes, which implies a devaluation of schoolwork and lower school engagement resulting in lower academic achievement (Francis 1999). Characteristics conducive to academic performance, such as zeal, the ability to cooperate and academic effort, are stereotyped as feminine (Heyder and Kessels 2015). Boys typically will not engage in stereotypical female behaviour, as this might compromise their masculine identity. Hence, boys are more likely to behave in a way that facilitates their construction of masculine identity, with lower educational performance as a consequence.[2] Scholars have also argued that behaviour that is assumed to be more feminine is more consistent with the school setting and therefore rewarded. Behaviour that is assumed to be more masculine, on the other hand, is incompatible with this school setting and therefore sanctioned (Heyder and Kessels 2013). The construction of masculinity, and its effect on

[2] Boys can accomplish a masculine identity with the use of different strategies and practices. Hence, the construction of a masculine identity might not necessarily reduce academic performance, if boys display masculinity in other contexts (e.g., sports) (Morris 2012).

academic performance, appears to be especially important around the mid-teenage years (Frosh, Phoenix, and Pattman 2002; Swain 2005). This suggests that when educational choices are made later in the educational career, boys will show lower academic effort affecting their educational performance and the expectations about their educational potential. Educational choices in late-tracking systems compared to early-tracking systems will then likely benefit girls' educational outcomes relative to boys'.

The gender gap in educational achievement is not constant over the educational careers of boys and girls. This implies that the sorting structure in tracks will affect boys and girls differently. In a system with early tracking, for example, at age 10, the academic performance of boys and girls is still relatively similar, and so is their mix of predictors of performance including non-cognitive skills and gender-typical behaviours. Tracking will have more negative consequences for boys when they must enrol in an educational track at a time when they lag behind girls in non-cognitive skills, show more behavioural problems, devalue educational achievement and have temporarily lower educational performance. Educational choices made at a later age during secondary education will therefore be beneficial for girls' educational attainment. Since we know that tracking decisions have long-term consequences (Dustmann 2004), tracking is likely to affect the gender gap in final educational attainment. Our core hypothesis is therefore that a higher age of first selection will favour women's final educational attainment.

Alternative explanations

Although we focus on tracking as the main contextual explanation, other important factors may also affect gender differences in education. In this section, we discuss three alternative explanations for the changing gender gap in educational attainment over time and cross-country differences: (1) the vocational orientation of educational systems, (2) gender-ideological context, and (3) the structure of the labour market. These factors are not only important potential predictors of the gender gap in attainment but also likely to be correlated with the tracking age.

First, educational systems not only vary in the extent to which they track students, but also in the extent to which they offer occupation-specific programmes. This vocational orientation of the educational system indicates the extent to which students are provided with job-specific skills as opposed to more general skills. An important argument is that acquiring work-related skills and having work-related experiences in a vocational programme can engage students who are at risk of dropping out of school. Prior research has indeed indicated that vocational education decreases dropout rates (Castellano, Stringfield, and Stone 2003), specifically when these vocational programmes do not begin very late (Pauly, Kopp, and Haimson 1994). Since

vocational education is a more appealing form of education to boys than to girls (Brunello and Checci 2007), we should consider that vocational education might be specifically beneficial for boys' educational attainment. Since the tracking age and vocational orientation are correlated (early tracked systems tend to have a larger vocational system, Bol and Van de Werfhorst 2013), it is especially important to take this system characteristic into account as well.

A second factor that can potentially influence the gender gap in education is gender-egalitarian attitudes within a country. This gender-ideological context is argued to influence gender differences in choices and preferences in education, for instance, by shaping expectations about future education and employment (Charles and Bradley 2002). Previous research has shown that a more gender-egalitarian context positively influences women's educational expectations, enrolment in tertiary education and final educational attainment (Buchmann, DiPrete, and McDaniel 2008; McDaniel 2010). Several studies have also demonstrated that although large country differences in the gender-ideological context exist, in all Western countries, there was a trend towards more gender-egalitarianism, especially during the second half of the 20th century (Brewster and Padavic 2000; Inglehart and Norris 2003). Therefore, changes in gender-egalitarian attitudes potentially explain between-country and over-time differences in the educational gender gap. Moreover, the gender climate is possibly related to the tracking age. Reforms to adjust the tracking age do not arise exogenously but are possibly embedded in wider cultural and political shifts towards equality. We therefore need to control for the gender-egalitarian climate.

A final confounder is the labour market demand. If individuals' anticipated future opportunities shape their present choices and performance (Goldin, Katz, and Kuziemko 2006), individuals are more likely to invest in human capital in contexts where they anticipate more opportunities to apply their skills on the labour market (Becker 1991; Charles and Bradley 2002). Therefore, changes in the labour market structure may affect gender differences in education. During the 20th century, the occupational structure changed rapidly (Castells and Aoyama 1994), experiencing a transition from manufacturing employment to service-related employment. The expansion of the service sector has been shown to be highly correlated with the growth of women's labour force participation (Becker, Hubbard, and Murphy 2010; OECD 2005) and can partly explain the rising female employment rates (Akbulut 2011). Hence, due to the increase in service-related jobs, more opportunities on the job market arose for women, making it more rational and attractive for them to invest in human capital. We thus expect that the size of the tertiary sector in a country and the structural changes in the labour market over time affect women's educational attainment. The proportion of service-sector employment may

furthermore be correlated to the tracking age, as both are potentially driven by broader societal trends towards women's empowerment.

2.3 Data and methods

Data

The individual-level data for this study are derived from the European Social Survey (ESS). The ESS is a cross-national dataset collected in 36 European countries. To date, the ESS has gathered 7 rounds of data, collected bi-annually between 2002 and 2014. In this study, we use all 7 rounds of the ESS for 21 European countries[3] for which we have information about educational institutional structures.[4] Information on contextual-level indicators, all time-varying, is gathered from various other sources described below.

The individual-level data are linked to contextual-level data on the basis of country-specific cohorts.[5] First, respondents are grouped into 5-year birth cohorts (e.g., 1961-1965, 1966-1970) within their countries. We match contextual-level variables to the cohorts in which the respondents were 20 years of age. We opt for age 20 because contextual features during adolescence might be specifically important for the decision to continue studying or to choose another path.[6] An example of our data structure is as follows: someone from Germany born in 1964 is grouped into the German 1961-1965 cohort and receives information on all contextual variables, e.g., average GDP per capita, for the period 1981-1985 in Germany, when he or she was 20 years of age. Information on the tracking age is not connected to 20 years of age but rather to the actual tracking age the respondents experienced during secondary education.

Based on the information we have on the tracking age; we can construct 13 birth cohorts (1921-1925 to 1981-1985). Country-cohorts containing less than 50 respondents are excluded from the sample. The sample size in our first analysis, including only the native-born population above 25-years of age, is N=143,883. This sample contains 21 European countries all including a maximum of 13 birth cohorts, resulting in 270 country cohorts.

[3] Countries included in this study are Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Great Britain, Greece, Hungary, Ireland, Italy, Latvia, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, and Sweden.

[4] Since Spain, with a tracking age of 6, was an extreme outlier, it was excluded from the sample. The results are highly similar when we include Spain. The selection at age 6 involved a primary school type specifically designed for farmer families.

[5] Country/year would be preferred; however, several contextual-level measures are not measured every year or are sometimes missing within a specific year. Moreover, several country/year combinations did not contain enough respondents, resulting in insufficient variation. Therefore, the data were aggregated to the cohort level.

[6] Moreover, choosing a younger age would result in losing many cohorts in the analysis. However, robustness checks are employed and showed that comparable results are obtained when opting for the age of 15 instead of 20.

We construct two other samples to examine the impact of other contextual features. Due to missing data on contextual-level control variables (depending on the variable we have data on these indicators since 1960 or 1970), we exclude birth cohorts before 1946-1950 from our second sample. The second sample contains 20 European countries all including a maximum of 8 birth cohorts (1946-1950 to 1981-1985), resulting in 124 country cohorts containing N=78,921 respondents.[7] In the last step of our analysis, we also include a measure of vocational education. This last sample contains 20 European countries all including a maximum of 6 birth cohorts (1956-1960 to 1981-1985), resulting in 92 country-cohorts containing N=52,612 respondents. Although these samples include a lower number of observations and cohorts, they allow us to test for competing explanations.

Individual-level variables

Educational attainment, our dependent variable, is operationalized as years of completed education.[8] Years of education is self-defined and measures the number of years of education a person has completed (including compulsory education), whether part-time or full-time, in full-time equivalents. Years of education is bottom-coded at 6 years to reflect that everyone has attained some primary education and top-coded at 25 to avoid strong leverage from outliers in educational attainment. A descriptive graph showing how men's and women's years of education evolved over time per country can be found in Appendix A, Figure A.1. The main individual-level independent variable of this study is gender (male=1, female=0).

Contextual-level variables

The tracking age is measured as the age at which students are initially streamed into separate educational trajectories. Information on the tracking age is gathered from the educational reform dataset constructed by Braga et al. (2013), which contains information on various reforms that influenced the educational systems in several European countries between 1929 and 2000. First, this measure is aggregated from a country/year-level measure to a country/cohort-level measure. Subsequently, birth cohorts are linked to the (average) tracking age that cohort experienced during secondary education. Countries that have an undifferentiated system are assigned the age at which students leave secondary education, as this is the moment students will separate into different (educational) trajectories. This measure

[7] Norway is not included in the KLEM database; therefore, Norway is excluded from the second and third samples.
[8] In the data, we could also operationalize educational attainment as the level of education (based on ISCED). In this study, we chose completed years of education, as we preferred a continuous variable that is easy to compare across countries and time. However, we also performed the analysis using education level, and the results were highly comparable, suggesting that our results are not sensitive to the measurement of the dependent variable. These results are presented in Appendix C, Table C.1.

is continuous, ranging from 10 to 16. For ease of interpretation, we subtract 10 from the tracking age so that 0 refers to the earliest tracking age.

The tracking age is arguably the most important feature that determines the level of tracking in an educational system. There are additional features, such as the number and length of different tracks. Due to these features' interrelatedness, it is highly likely that when the tracking age changed within an educational system, these other characteristics changed as well. However, theoretically, we argue that specifically the age of first selection interacts with gender differences in development and is therefore the main characteristic relevant for gender differences in education. Additionally, age of first selection is empirically shown to be the main feature explaining inequality of opportunity (Horn 2009). Moreover, Pekkarinen (2008) showed for Finland that change in the tracking age—not change in the curriculum—influenced gender differences.

The vocational orientation of the educational system is operationalized as the percentage of students enrolled in vocational programmes at public and private secondary education institutions relative to students enrolled in all secondary education programmes. Vocational education is described as education that is designed for students to acquire knowledge, skills and competencies specific to a particular occupation, trade or class of occupations, which may involve work-based components. Data are gathered from the UNESCO Institute for Statistics, which contains country-specific information on this measure from 1970 onwards. This measure is aggregated from a country/year-level measure to a country/cohort-level measure and connected to the time when birth cohorts were 15 years of age.[9] The correlation between vocational orientation and the tracking age in our data is -.30. This negative association suggests that in countries that track earlier, the percentage of students in secondary school who are enrolled in vocational programmes is higher.

The gender-ideological climate is operationalized using items from all rounds of the European Values Study (EVS, 1981-2008) and World Values Survey (WVS, 1981-2014). We generate a gender attitude scale based on the extent to which respondents agree with eight statements: (1) A woman has to have children to be fulfilled. (2) A child needs a home with a father and a mother. (3) A woman can be a single parent. (4) A working mother can establish as warm a relationship with children as a non-working mother. (5) Being a housewife is as fulfilling as having a paid job. (6) A pre-school child suffers by having a working mother. (7) Women truly want a home and children. (8) When jobs are scarce, men should have a greater right to a job than women. These items are selected on the basis of broad availability with respect to countries and years, as well as their use by previous research (Davis and Greenstein

[9] Vocational education is linked to respondents at age 15 instead of those at age 20, as at age 15, vocational orientation can influence people's educational choices: whether to enter general or vocational education.

2009; McDaniel 2008). The items are recoded so that a higher score indicates a more gender-egalitarian attitude and subsequently standardized. A principal component analysis shows that all items load above 0.4 on one factor, and the Cronbach's alpha of these items is 0.65. We take the mean of the eight items to construct the gender attitudes scale. Respondents who answered less than four of the eight questions are coded as missing. However, the attitudes that we derive from the EVS and WVS are for the full population, not for specific cohorts. For this reason, we use a regression model to predict the gender attitudes by birth-cohort, age, country and interactions among all three. Subsequently, we use marginal effects to predict the average attitudes of young adults (21- to 30-year-old population) within a country and a cohort between 1941 and 2005. On this final measure, a higher score indicates a more gender-egalitarian climate. This measure is matched to individual-level data when birth cohorts were 20 years of age.

The labour market demand is operationalized as the percentage of total hours worked by employees in the tertiary sector. The data are gathered from the EU KLEMS database (November 2009 release).[10] This database contains internationally comparable industry-level indicators of input, output and productivity for 30 countries – mostly European Union member states, but also the U.S., Australia, Japan and Korea – from 1970 up to 2007 (O'Mahony and Timmer 2009). This dataset includes information about total hours worked by employees in the different industries per year and per country. The percentage of hours worked by employees in the tertiary sector relative to hours worked by employees in all sectors is calculated.[11] This measure is then aggregated from a country/year-level measure to a country/cohort-level measure and subsequently linked to individual-level data when birth cohorts were 20 years of age.

An overview of all contextual-level indicators, including a short description, years of availability and the source, is presented in Table 2.1.

[10] The data are publicly available on http://www.euklems.net. The construction of this database is funded by the European Commission, Research Directorate General as part of the 6th Framework Programme, Priority 8, "Policy Support and Anticipating Scientific and Technological Needs" and as a part of the 7th Framework Programme, Theme 8: Socio-Economic Sciences and Humanities. 24 research institutes and national statistical institutes participated in compiling the EU-KLEMS database (O'Mahony and Timmer 2009, see Appendix A).

[11] The tertiary sector includes employees who work in wholesale and retail trade, hotels and restaurants, transport, storage and communication, finance, insurance, real estate and business services, public administration and defense, education, health and social work, private households, extra-territorial organizations and bodies, and other community, social and personal services.

Table 2.1. Description, availability and sources of contextual-level variables

Contextual-level variables	Short description	Years of availability	Birth-cohorts availability	Source
Tracking age	The age of selection into separate educational trajectories.	1929 – 2000	1921-1925 to 1981-1985	Braga, Checchi, and Meschi (2013)
Vocational enrolment	The percentage of students in secondary education that are enrolled in vocational programmes.	1971 onwards	1956-1960 to 1981-1985	UNESCO Institute for Statistics
Gender-ideological climate	The average gender-egalitarian attitudes of young adults.	1941 – 2005	1946-1950 to 1981-1985 [a]	World Values Survey and European Values Study
Tertiary sector [b]	The percentage of total hours worked by employees in the tertiary sector.	1970 onwards	1946-1950 to 1981-1985	EU KLEMS database
Log(GDP)	The logarithm of GDP per capita in US dollars.	1960 onwards	1946-1950 to 1981-1985 [a]	World Data Bank

[a] These variable could be connected to earlier birth-cohorts, however in the analysis we only use them in combination with other country-level variables. Therefore, we only provide descriptive statistics from the 1946-1950 cohort onwards.
[b] Not available for Norway

Control variables

We include GDP as a control variable on the country cohort level. This variable is measured as the GDP per capita in US dollars, and data were collected from the World Data Bank, which contains country-specific information on this measure from 1960 onwards. In addition, the logarithm is taken to control for the skewness of the variable. As with all contextual variables besides the tracking age, the measure is matched to individual-level data for birth cohorts that were 20 years of age.

At the individual level, we include the father's and the mother's years of education. For both measures, we convert International Standard Classification of Education (ISCED) levels into years of schooling: 6 years for ISCED=0-1 (less than lower secondary education completed), 9 years for ISCED=2 (lower secondary education completed), 12 years for ISCED=3 (upper secondary education completed), 14 years for ISCED=4 (post-secondary non-tertiary education completed), 16 years for ISCED=5-6 (tertiary education completed). We also control for the father's occupation when the respondent was fourteen years old by including five out of the six following dummy variables: whether the father worked in (1) Routine manual and service occupations, (2) Semi-routine/manual/service occupations, (3) Technical and craft occupations, (4) Clerical and intermediate occupations/middle managers, (5) Higher

administrator occupations, or (6) Professional and technical occupations. The descriptive statistics of all individual- and country/cohort-level variables are presented in Table 2.2.

Table 2.2. Descriptive statistics

	Mean	SD	Min	Max	N
Individual-level variables					
Years of education	12.683	3.877	6	25	143,883
Male	.469		0	1	143,883
Father's years of education	9.892	3.438	6	16	143,883
Mother's years of education	9.284	3.129	6	16	143,883
Father's occupation					
Routine manual and service	.243		0	1	143,883
Semi-routine/manual/service	.175		0	1	143,883
Technical and craft	.284		0	1	143,883
Clerical and intermediate	.150		0	1	143,883
Higher administrator	.051		0	1	143,883
Professional and technical	.097		0	1	143,883
Contextual-level variables					
Tracking age (10 years=0)	3.155	1.976	0	6	143,883
Gender-ideological climate	.120	.217	-.300	.616	78,921
Tertiary sector	59.377	7.690	40.056	77.419	78,921
Vocational enrolment	28.623	13.100	2.129	53.734	52,612
Log(GDP)	9.006	.899	6.599	10.595	78,921

Methods

We employ a two-level multilevel model to examine the extent to which the effect of gender on years of education is moderated by the age of first selection. This model takes into account that individuals are nested within country-cohorts. We include country fixed effects, which allows us to control for all time-invariant country-specific factors and indicates that the estimators of this model are not contaminated with spurious effects of stable, unmeasured country characteristics (Verbeek 2004). We also incorporate cohort fixed effects, which enables us to control for unobserved general time trends that affect all countries equally. In this model, we thus exploit the within-country over-time variation in tracking age to determine whether policy changes affected the gender slope on years of education. The general equation for this model is the following:

$$Y_{ij=}\beta_0 + \sum_{l=l-1} \beta_x L_j + \sum_{k=k-1} \beta_k K_j + \beta_3 gender_{ij} + \beta_4 trackage_j + \beta_5 gender_{ij} * trackage_j + \beta_6 X_{ij} + u_{0j} + u_{1j} gender_{ij} + e_{ij}$$

In this equation, Y_{ij} is the number of years of completed education for individual *i* in country cohort *j*; β_x estimates the fixed effects for countries by adding dummies for country L; β_k estimates the fixed effects for cohorts by adding dummies for cohort K; β_3 to β_5 are the estimates of the main effects and cross-level interaction between gender and the tracking age; β_6 estimates the effects of all individual-level control variables. u_{0j} is the error term at the country cohort level; $u_{1j}gender_{ij}$ is the error term of the slope variation in gender, as this model includes a random slope for gender; e_{ij} is the error term at the individual level.

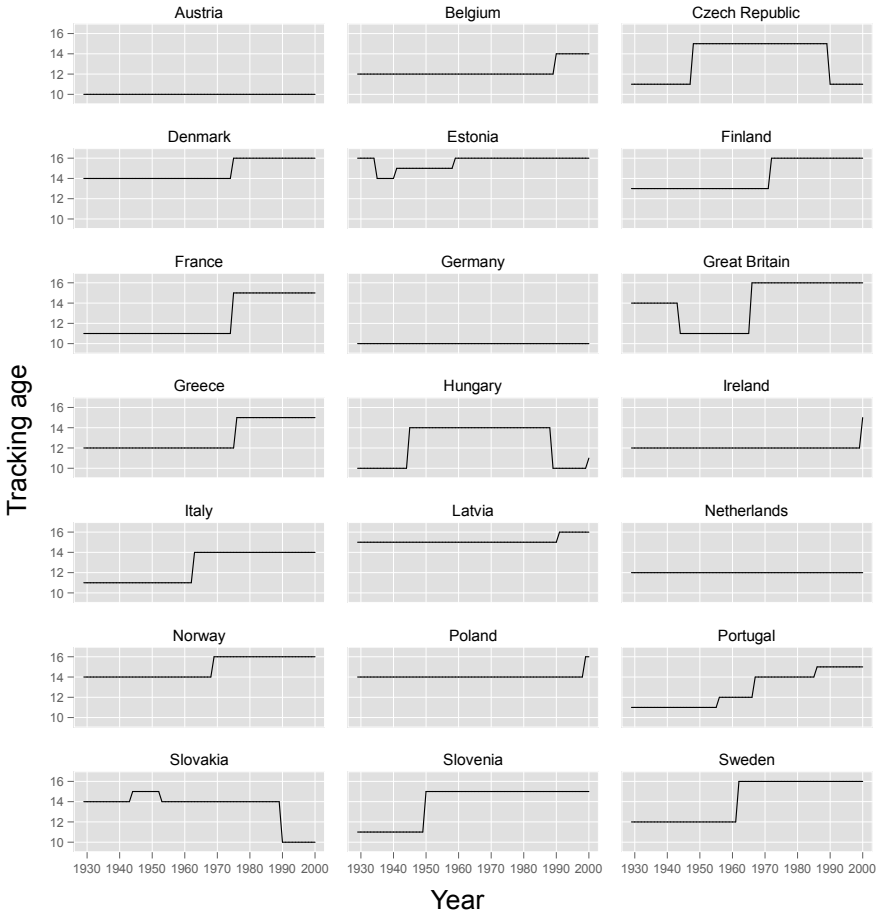
To test whether the results are sensitive to different specifications of the random effects structure, we also performed the analysis using two other multilevel modelling approaches: a cross-classified model and a hybrid model (Schmidt-Catran 2016; Schmidt-Catran and Fairbrother 2016). The general equations and explanations of these methods can be found in Appendix B.

2.4 Results

Descriptive results

This paper began with the argument that even though the relative educational attainment of women is increasing in almost all Western countries, large variation exists across countries and over time. Proof for this argument is demonstrated in Appendix A, Figure A.1. Before we continue to our main analysis, we look briefly at the extent to which the tracking age actually differs between countries and changes over time. Figure 2.1 displays the variation in the tracking age over time for all countries included in our sample. It shows that large differences in the tracking age not only exist between countries, with some countries selecting students much earlier than other countries, but also sometimes changed within countries. Only in 3 of the 21 countries does the tracking age remain stable (Austria, Germany and the Netherlands). In general, this figure displays an upward trend in the age of first selection in almost all countries, indicating that most countries postponed the tracking age. According to our argument, this upward trend might have contributed to the "rise of women". Figure A.2, Appendix A, presents descriptive statistics per country of all other contextual-level variables.

Figure 2.1. Variation in the tracking age over time for all countries

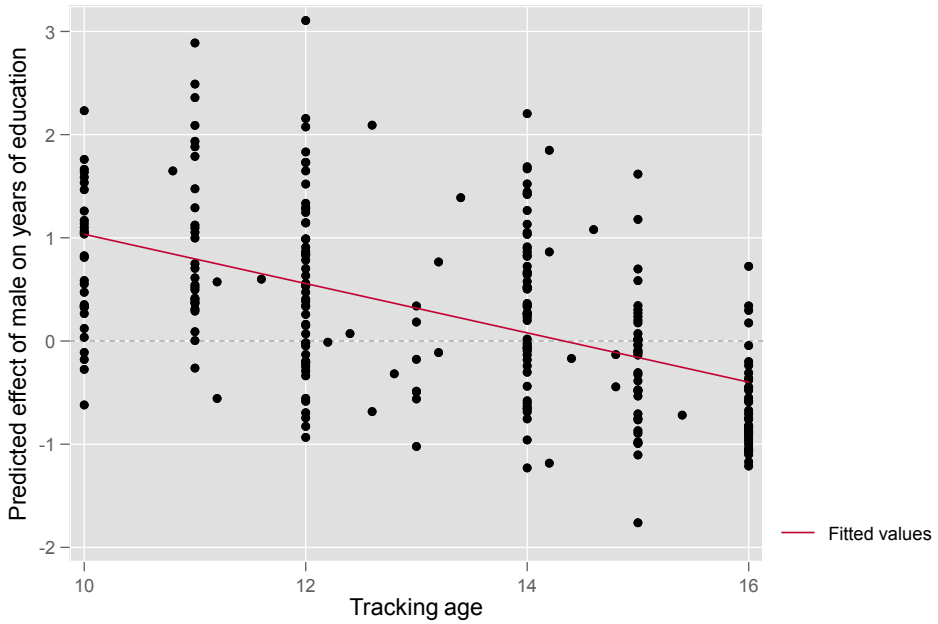


Source: Braga et al. (2013)

Moreover, we first descriptively examine whether the relationship between gender and education appears to be different for societies with different tracking ages. This relationship is presented in Figure 2.2, which displays the effects of being male on completed years of education for different ages of selection estimated by separate OLS regression models for each country cohort.[12] Each dot in the scatterplot is an estimated effect of being male (relative to being female) on years of education for one country cohort.

[12] These models controlled for parents' education and the father's occupation.

Figure 2.2. Predicted effect of male on completed years of education for different ages of selection



Note: Each dot in the scatterplot is estimated by separate OLS regression models and displays effect of male (relative to female) on completed years of education for one country-cohort. Equation: Y(years of education) = b0 + b1*male + b2*X. In the equation, b2 estimates the effect of all individual-level control variables, which are parent's years of education and dummies for father's occupation.

The figure clearly shows that whereas the effect of being male on years of education is mainly positive when the tracking age is 10, it decreases and even becomes mostly negative when the tracking age is higher. This is in line with our expectations that a higher tracking age benefits girls' educational attainment. The relationship displayed in the figure may, however, be due to extraneous effects such as general time trends; in general, there was an upward trend in the number of years of education women completed and in the age of first selection. In our main analyses, we will control for the general time trend and investigate if a policy change (change in the tracking age) influenced the effect of gender on years of education.

Multivariate results

The baseline model of the two-level multilevel model including only country and cohort fixed effects shows that even though these fixed effects account for a large part of the contextual-level variance, 2.4% of the total variance is at the country/cohort-level. After including individual-level variables, we test the significance of a random slope for gender. A log likelihood ratio test shows that the model fit is significantly improved when this random slope

is included ($\chi 2(2)=1529.72$, $p<0.0001$), indicating that the effect of gender on education varies across country-cohorts.[13]

Table 2.3 displays the results of the two-level multilevel model of years of education. This model includes country and cohort fixed effects, individual-level indicators, tracking age and the cross-level interaction between tracking age and gender to potentially explain the random slope of gender.[14] The individual-level indicators, father's and mother's years of education and father's occupation, all have a significant and positive effect on years of education.

The positive and significant main effect of gender indicates that on average, men complete almost one more year of education than girls when the tracking age is 10. We also find a positive main effect of the tracking age, with a one-year increase in the tracking age leading to a quarter-year more education for women. The cross-level interaction between the tracking age and gender shows that the gender effect on years of education differs significantly by the age of selection. In line with our expectations, this effect is negative, indicating that the advantage of being male weakens as the tracking age increases. This model predicts that the effect of being male on years of education even reverses when the tracking age surpasses 14 years. With a 0.25 change in the gender effect on years of education for each year that tracking is delayed, this effect is substantial. It is important to note that this effect is net of the general time trend that we observe across these countries. This cross-level interaction explains 33.9% of the random slope for gender.

The results of the cross-classified model and the hybrid model of years of education can be observed in Appendix B. The estimates are consistent across the different models: for every one-year increase in the tracking age, the effect of being male on years of education decreases by a quarter of a year, which indicates that our results are robust to different model specifications.

In Table 2.4, we present the results for when we control this finding for other contextual factors. Since a specific higher-level variable—vocational education—is only available for later cohorts, two separate two-level multilevel models with country and cohort fixed effects are estimated: one without this parameter (Model 3) and one with this parameter (Model 5). Since the sample size changes with the inclusion of the contextual control variables, the differences in effects might be due to the different sample size as well. To ensure that changes in the results are not caused by the different sample size but rather by the inclusion of other factors, we repeated Model 1 with the new sample size restriction in both models (Model 2 and Model 4).

[13] The change in degrees of freedom is 2 because we also allow the intercept and slope to covary.
[14] Including country-specific linear time trends or even country-specific linear time trends in the gender gap instead of a general time trend did not alter our results.

Table 2.3. Multilevel linear regression models of years of education, with a random slope for male and country and cohort fixed effects

	Model 1	
	B	SE
Individual-level variables		
Father's years of education	.214 ***	.004
Mother's years of education	.212 ***	.004
Father's occupation		
Routine manual and service	ref.	
Semi-routine/manual/service	.172 ***	.025
Technical and craft	.462 ***	.023
Clerical and intermediate	1.320 ***	.028
Higher administrator	1.583 ***	.043
Professional and technical	1.545 ***	.036
Male	.992 ***	.084
Contextual-level variables		
Tracking age	.241 ***	.034
Cross-level interaction		
Tracking age * male	-.235 ***	.022
Constant	4.893 ***	.210
Country fixed effects	Yes	
Cohort fixed effects	Yes	
Variance components		
Country-cohort	.406 ***	
Individuals	9.201 ***	
Random slope male		
Country-cohort	.413 ***	
N(observations)	143,883	
N(country-cohorts)	270	

Two-tailed test: * p<.05. ** p < .01. *** p < .001.

Model 3 shows that after controlling for potential confounders at the contextual level, we still observe a positive significant main effect of gender and the tracking age and a negative significant cross-level interaction, although the effect size is smaller. We do not find evidence that gender-ideological context during adolescence influences people's years of education or that the effect differs between men and women. The main effect of tertiary sector is not

significant but the cross-level interaction with gender is negative and significant, indicating that the positive effect of being male on years of education weakens when the size of the tertiary sector grows. The contextual-level control variable GDP shows that as GDP increases, years of education increases. The significance and direction of the effects of individual-level control variables remain stable.

In Model 5 (Table 2.4), we also include the size of the vocational sector of the educational system. The results found in this model are slightly different from the results found in Model 3, but this is mainly due to the lower number of country-cohorts included in this model and not to the addition of the new indicator. Due to the loss of earlier cohorts, the effect of gender is not significant (see Model 4), but the outcomes for the tracking age and the corresponding cross-level interaction remained the same: a positive effect of the tracking age on years of education that is significantly weaker for men than for women. Again, we find no evidence of an effect of gender egalitarian context on (gender differences in) years of education. Moreover, in this model, we find that a larger tertiary sector is associated with an increase in years of education, but this pattern is not significantly different for men and women. For vocational enrolment, we find a negative significant coefficient, indicating that vocational enrolment is negatively associated with the average years of education. For every standard deviation increase in vocational enrolment, completed years of education decreases by 0.22 years (which is equal to 0.06 standard deviations in years of education). The non-significant cross-level interaction illustrates that this effect does not vary for men and women.

In sum, irrespective of the modelling technique that we use or the control variables that we add, we find that the gender gap in years of education is associated with the tracking age. When societies postpone the moment of tracking, female students benefit more from it. To obtain more insight in the size of this effect, Figure 2.3 displays the marginal effect of being male on years of education for different ages of selection based on our final model (Model 5). This figure demonstrates that when a country raises the tracking age, the model predicts that gender differences in completed years of education in favour of women also increases. It also shows that when children are tracked at an age of 10, 11 or 12, we do not predict a (significant) difference in completed years of education between boys and girls. However, when the tracking age is 16, girls outperform boys in education, with a difference of approximately 0.15 of a standard deviation on years of education. This is a large effect for a system-level factor such as tracking, indicating the importance of a contextual explanation of the gender gap in educational attainment.

Table 2.4. Multilevel linear regression models of years of education, controlling for alternative explanations

	Model 2		Model 3		Model 4		Model 5	
	B	SE	B	SE	B	SE	B	SE
Individual-level variables								
Father's years of education	.186 ***	.005	.186 ***	.005	.159 ***	.006	.159 ***	.006
Mother's years of education	.205 ***	.005	.205 ***	.005	.210 ***	.006	.210 ***	.006
Father's occupation								
Routine manual and service	ref.		ref.		ref.		ref.	
Semi-routine/manual/service	.087 *	.037	.087 *	.037	.031	.045	.031	.045
Technical and craft	.383 ***	.034	.382 ***	.034	.329 ***	.042	.327 ***	.042
Clerical and intermediate	1.197 ***	.039	1.196 ***	.039	1.023 ***	.047	1.021 ***	.047
Higher administrator	1.395 ***	.057	1.395 ***	.057	1.153 ***	.068	1.151 ***	.068
Professional and technical	1.446 ***	.049	1.446 ***	.049	1.342 ***	.059	1.340 ***	.059
Male	.505 ***	.097	.420 ***	.096	.183	.110	.210	.114
Contextual-level variables								
Tracking age	.178 ***	.042	.208 ***	.041	.137 **	.047	.157 ***	.039
Gender-ideological climate [a]			-.218	.547			.739	.457
Tertiary sector [a]			.012	.020			.058 **	.017
Log(GDP)			1.232 ***	.239			1.034 ***	.196
Vocational enrolment [a]							-.017 **	.005

Two-tailed test: * p<.05. ** p < .01. *** p < .001.

[a] For ease of interpretation, these variables have been mean centered in the models.

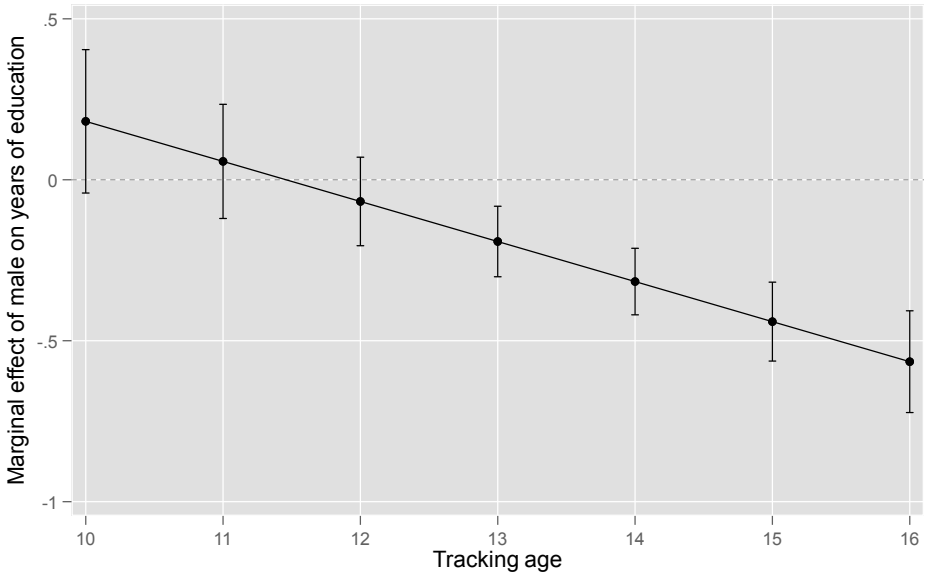
Table 2.4. Multilevel linear regression models of years of education, controlling for alternative explanations (continued)

	Model 2		Model 3		Model 4		Model 5	
	B	SE	B	SE	B	SE	B	SE
Cross-level interactions								
Tracking age * male	-.180 ***	.025	-.147 ***	.026	-.125 ***	.026	-.124 ***	.027
Gender climate [a] * male		.317	-.027	.317			-.035	.331
Tertiary sector [a] * male		.009	-.025 **	.009			-.008	.010
Vocational enrolment [a] * male							-.001	.004
Constant	7.153 ***	.187	-2.276	1.769	8.343 ***	.172	-.634	1.772
Country fixed effects	Yes		Yes		Yes		Yes	
Cohort fixed effects	Yes		Yes		Yes		Yes	
Variance components								
Country-cohort	.317 ***		.280 ***		.201 ***		.122 ***	
Individuals	9.663 ***		9.663 ***		9.387 ***		9.387 ***	
Random slope for male								
Country-cohort	.273 ***		.235 ***		.178 ***		.172 ***	
N(individuals)	78,921		78,921		52,612		52,612	
N(country-cohorts)	124		124		92		92	

Two-tailed test: * p<.05. ** p < .01. *** p < .001.

[a] For ease of interpretation, these variables have been mean centered in the models.

Figure 2.3. Marginal effect of male on years of education by the tracking age



Note: This figure is based on model 5, Table 2.4. The 95 percent confidence intervals are presented.

Sensitivity analysis

To assess the reliability of our outcomes and check for potential influential countries, we perform sensitivity analyses excluding specific clusters of countries. Table 2.5 displays four models excluding Scandinavian countries, Eastern European countries, Southern European countries, and West European countries, respectively. In these models, we control for other contextual factors, although the coefficients are not displayed in the table.[15] The results presented in Table 2.5 are highly comparable to those for the full sample. The main effect of being male is overall positive and significant, except for the model excluding Western European countries—indicating boys' presumably much higher educational attainment over girls in these countries. In all models, we still find a positive and significant main effect of the tracking age and a negative and significant cross-level interaction, indicating that excluding specific clusters of countries did not dramatically influence our outcomes. While all models show the same pattern, the effects that we find are of course affected by the inclusion (or exclusion) of countries. Nevertheless, the general pattern is robust to selecting different sets of countries.

[15] Only vocational enrolment is not included in these models due to the limited number of cohorts covered. However, including vocational enrolment in these models did not change the results substantially.

Table 2.5. Multilevel linear regression models of years of education, testing for influential countries

	W/o Scandinavian countries		W/o Eastern European countries		W/o Southern European countries		W/o Western European countries	
	B	SE	B	SE	B	SE	B	SE
Male	.434 ***	.098	.447 ***	.100	.369 ***	.092	.148	.248
Tracking age	.219 ***	.043	.210 ***	.048	.096 **	.032	.183 *	.077
Tracking age * male	-.127 ***	.029	-.129 ***	.028	-.166 ***	.026	-.112 *	.054
N(individuals)	62475		68530		67249		38509	
N(country-cohorts)	100		103		101		68	

Note: Scandinavian countries: Denmark, Sweden, Finland and Norway. Eastern European countries: Czech Republic, Estonia, Hungary, Latvia, Poland, Slovakia and Slovenia. Southern European countries: Greece, Italy and Portugal. West European countries: Austria, Belgium, France, Germany, Great Britain, Ireland and the Netherlands. Based on Model 3, Table 2.4.

Two-tailed test: * p<.05. ** p < .01. *** p < .001.

2.5 Conclusion and discussion

Although it is well known by now that women have surpassed men in educational attainment, a good understanding of the reverse educational gender gap remains lacking. In this study, we have focused on the extent to which the educational institutional structure, and more specifically, the tracking age, is associated with gender differences in completed years of education. This study is the first to systematically investigate this relationship for a large number of European countries over a long time period. Moreover, we examined various rival dynamic theoretical explanations that potentially affected the reversal of the gender gap.

The main finding of our study is that later tracking improves women's completed years of education relative to men's. This result is robust to different model specifications and tests of several rival explanations. Although this study shows that the postponement of tracking favours women's final years of education, the theoretical explanation underlying this relationship requires further study. We have offered two interrelated mechanisms. First, several scholars have argued that the gender gap in non-cognitive skills increases during secondary education, indicating that the gender gap in educational achievement widens during this period as well. Second, studies point to the importance of gendered identities during secondary education, which affect boys' and girls' performance in education. The general idea underlying these mechanisms is that there is an interaction between the timing of tracking and the varying differences between boys and girls across their school careers. In line with our results, these two mechanisms suggest that in particular, educational choices at the end of lower secondary education will favour girls' educational attainment.

According to Baker (2014), societies have become "schooled societies", embracing a cultural model in which the social positioning of individuals and the structure of society are strongly based on educational qualifications. This cultural model has affected the attraction of higher-level qualifications and may have contributed to a massive expansion of education especially among women (as they lagged behind men). Nevertheless, our study shows that the gender-specific trend in attainment is not the same in all societies, as it was much easier for women to catch up with men (and surpass them) in systems with later tracking. A neo-institutionalist perspective on our findings could, however, imply that reforms promoting later tracking fit schooled societies, as expansion is more easily achieved in later-selecting systems. Hence, a cultural (rather than functional) model of expansion could lie behind both reforms and rising attainments, which would not be at odds with our argument. Nevertheless, it would be difficult for a cultural model to explain the reversal of the gender gap beyond the equalization of attainment between men and women. Moreover, if our results were interpreted

from a neo-institutionalist perspective, it would mean that schooled societies have not evolved everywhere at the same speed.

Women are increasingly becoming more educated than men, and our study shows that educational systems contribute to this phenomenon. However, it is important to stress that there are still large gender gaps in favour of men in education and the labour market, and more years of schooling does not necessarily translate into more equality in these areas. Bradley (2000) showed that a higher representation of women in higher education did not translate into gender parity in the occupational structure, as gender segregation within the educational system remained high. The postponement of tracking is suggested to be associated with changes in the combination of fields of study represented at specific educational levels, for example by upgrading (e.g., from the secondary to the tertiary level) of traditionally female-labelled educational fields, such as nursing and teaching (Bradley and Charles 2004; Breen et al. 2010). This suggests that, even though women in late-tracking systems complete more years of schooling, women in both systems end up with the same occupational credentials. Nevertheless, research shows that girls are less likely to end up in female-typical educational fields in systems with later tracking, as early gendered career aspirations are less likely to be translated into final educational and occupational outcomes in late-tracking systems compared to early-tracking systems (Charles et al. 2001). Additionally, Pekkarinen (2008) showed that the gender wage gap decreased after the tracking age increased, suggesting that later tracking, potentially via educational attainment, also contributes to women's labour market success. To obtain a more comprehensive understanding of the impact of educational tracking, future research should examine whether tracking age also affects gender differences in fields of study and other educational and labour market outcomes.

As we have shown in this paper, educational systems change over time, which might result in different opportunities for different groups. Our findings indicate an interesting trade-off between gender equality and equality with respect to socioeconomic backgrounds in educational opportunity. Whereas a higher tracking age has repeatedly been shown to diminish socioeconomic inequalities in educational attainment (Van de Werfhorst and Mijs 2010), our study demonstrates that it enhances gender differences in education. Future research should investigate the above-mentioned trade-off more thoroughly and examine whether a combination of educational institutional structures can reduce both forms of educational inequality. In a similar vein, future research should investigate the intersection between gender and socio-economic background because the effect of tracking on boys' and girls' educational performance might differ across social classes.

In this article, we looked explicitly at between-school tracking. While not all countries track students between schools, most countries employ some type of ability sorting, for

example, within schools, such as in the United States or the United Kingdom (Chmielewski 2014; LeTendre, Hofer, and Shimizu 2003). Also in these countries students have to make curricular choices, e.g., content and level of difficulty (Lucas 2001). Our study argues that gender differences are not constant over the educational career, and that choices made at the end of lower secondary education are more likely to benefit girls, irrespective of whether it concerns between-school track choices or within-school curriculum choices. Future research should look at these systems and investigate if tracking has the same effects when it takes a less institutionalized form. Additionally, our findings potentially have implications for any intended reform, both in- and outside of Europe, targeting the educational selection and differentiation of specific age groups. For instance, debates in the United States about the consequences of a transition from middle to high school (e.g., Weiss and Bearman 2007; Pharris-Ciurej, Hirschman, and Willhoft 2012), may learn from the relevance of building or removing transitions at crucial age points for gender inequalities.

In conclusion, our article shows that the set-up of the educational system can influence the extent to which boys and girls perform well in education. Specifically, this study is the first to provide international evidence that female dominance in educational attainment is enhanced in educational systems that have postponed the tracking age. A higher age of first selection benefits women's completed years of education and therefore contributes to the gender gap in education, indicating that future studies on "the rise of women" should focus on contextual explanations and mechanisms for cross-national and cross-temporal differences. Such contextual explanations cannot neglect the educational institutional structure.