

Introduction

No child is alike. Each child has a distinct personality and has its own talents and skills. Teachers in school will therefore try to approach each child in a different way to ensure the child will be able to learn optimally and develops the necessary non-cognitive skills. In all education systems and in all school types some form of grouping students is done, to the very least by grouping students together based on age and grade. When within grades students are formally separated based on ability level, this is called tracking, or formal differentiation. Tracked systems often consist of students of different tracks to be placed into different building or even different schools. The different tracks in which students are placed have different curricula tailored to the ability levels of the student population. This adjusted curriculum, combined with a more ability homogenous class which should facilitate teaching, is thought to enhance learning outcomes.

Notwithstanding the mentioned theoretical benefits, the use of tracking in secondary education systems is a contested practice in western society and academia. Questions which are often raised are (1) Does tracking increase the average cognitive development? and following from that (2) Does tracking increase inequality among students? (OECD, 2010). Opponents argue that allocating the more able students to the high track, will leave the less able students without the more able peers to learn from. If this is the case, any gain achieved by the high ability students could be at the expense of the low ability students, increasing inequality in outcomes and perhaps even lowering average performance (Hanushek and Woessmann, 2006). Besides lower ability peers, students in the low track could also potentially lose out since they receive a less challenging curriculum, although it is designed for their ability level, and the lower ability school might receive less resources (in terms of facilities, or lower educated teachers). The second question also arises from the fact that specific groups of students, for instance students of lower parental background or the relatively young students in the class, are more likely to be allocated to the low track (e.g. Dustmann, 2004, Muehlenweg and Puhani, 2010). When track choice has long run consequences this might further disadvantage these students.

1.1 Aim of this thesis

This thesis provides evidence for the two questions asked above: Whether tracking has a positive effect on student learning and whether tracking increases inequality within education systems and societies. It thereby aims to provide the reader with a deeper understanding of tracking and its effects on student performance and inequality. To start, Chapter 2 provides direct evidence on the first question looking at the effect of tracking on student performance. The next two chapters look closer at the performance differences between tracks and tracking regimes. The aim of chapter 3 is to consider

how implementation of tracking by schools can influence the outcomes of tracking. It does so by looking at whether using prior performance to decide on track placement is related to better outcomes, both in student performance and inequality. Chapter 4 looks at whether the student on the margin of going to the low or high track is better off in terms of cognitive and non-cognitive outcomes in the high track. Finally, the second question is addressed in Chapters 5 and 6. These chapters look at the effects of tracking on two types of inequalities: those caused by parental background and those caused by relative age within the class.

1.2 Outline and results

Chapter 2 starts off by providing a deeper understanding of tracking by presenting evidence on whether tracking increases the average cognitive development of students. Previous studies have looked at this question before but this has led to conflicting results (e.g. Hanushek and Woessmann, 2006, Ariga and Brunello, 2007, Jakubowski, 2009, and Van Elk *et al.*, 2011). Analyzing the effect of tracking on student performance is problematic since all the variation is between countries and much less variation exists within countries. Less within country variation exists since tracking is often a national or statewide policy leaving no room for specific schools or regions to deviate from this policy.¹ If only cross country variation exists then any country differences will confound the analyses since it is impossible to disentangle the effects due to tracking from effects due to other country characteristics. To obtain unbiased estimates of the effect of tracking on educational outcomes, this endogeneity has to be taken into account either by longitudinal individual level data or by statistical methods to remove it. In Chapter 2 of this thesis one of the statistical methods to remove endogeneity is used. It uses an instrumental variable approach to be able to estimate the effect of tracking on student performance. The employed instrument is the political pressure by Napoleon in the 19th century. This instrument has a relation with the extent of tracking in different European countries, and the results show that tracking has a positive effect on student performance at age 15. However, previous research has successfully used related instruments for other institutions which cast doubts on the exclusion restriction and the causal interpretation of the results.

Chapter 3 of this thesis examines whether the implementation of tracking influences student performance and inequality. In a tracked system students need to be allocated over the different tracks. Since the main differences between tracks are on the content and level of the curriculum, it seems obvious that students are also allocated to the

1. Exceptions are variation between states/*Länder* or possibilities to postpone tracking in some tracked system. In Germany the *Gesamtschule* provides such an option and also in the Netherlands students can postpone strict tracking by attending bridge classes in the first or first two years of secondary education.

tracks based on ability. However in some countries school principals select students based on religious or artistic criteria, or parents have a large say in the track placement of their children (e.g. Dustmann, 2004, Barg, 2013). Chapter 3 of this thesis makes use of a question answered by school principals on whether they consider prior performance on accepting the student to the school. From the results it follows that in a system with four or five tracks, tracking is positively related to student performance if students are indeed selected based on performance. When school principals do not consider prior performance when accepting the student, students do not benefit from tracking. A further finding is that the influence of parents on student performance is lower when prior performance is always considered. This logically follows from the use of prior performance for track placement which often replaces parental preferences as the leading criterion for track placement.

One possible downside of tracking is that it could aid the students in the high track at the expense of the students in the lower tracks. Before looking into inequalities for specific groups, it is useful to study the underlying assumption on which this inequality is based: Individual students in the low track would perform better if they would be in the high track. Chapter 4 of this thesis studies this question by making use of the marginal student, the student who is on the border of being allowed or not to go to the high track. Looking at the marginal student allows us to look at the effect of being in the higher track without having the confounding factor of a changing ability composition, which is the case if a large number of students would move from the low to the high track. In the Netherlands, at the time of this research, secondary schools decide on track placement based on an elementary school exit test and a track recommendation of the elementary school teacher. Using these two assignment variables we are able to use the threshold for when students are allowed to go to the high track as a discontinuity in the assignment. For the marginal student, being placed in the high track improves reading and IQ scores and the perceived probability to finish the degree. These improved outcomes are reached solely from being in the high track, without changing the ability composition of the higher track. The characteristics of the high track, consisting of better able peers, a different curriculum, and more, cause the marginal student in the high track to do better than if the student would have been in the low track. Being in the high track does not improve mathematics scores or personality traits, as the Big Five or motivation. Although the methodology only allows us to look at the marginal student, for this student and the student's reading and IQ scores and the perceived probability to finish the degree it does matter whether the student is placed in the low or high track.

The last two chapters finish the analyses of tracking and its consequences for inequality by looking at the effects of parental background (Chapter 5) and at whether tracking influences the effects of relative age (Chapter 6). If track placement is done

on prior performance students in the low track will be students with lower ability. However, often students in the low track are also students of lower parental background and the relatively younger students. Students with lower parental background perform worse in school due to, among other things, worse health, nutrition, parental support, and social capital (see for an overview Jackson, 2013). As a consequence of this lower performance they are more often placed into the low track. But students of lower parental background are also less likely to go the high track since less is expected from them by teachers or teacher respond differently to students with higher educated parents (e.g. Lavy and Sand, 2015, Barg, 2013, Jussim and Harber, 2005). That students with lower parental background have lower educational outcomes before and after tracking has taken place, and are more likely to go to the low track is well documented. The innovation of Chapter 5 is that it looks at whether there is an additional effect of parental background on track placement over and above the influence of parental background on performance in general. If this is the case, then higher educated parents are able to enlarge the advantage their children have in a tracked system as compared to in a comprehensive system. The research question is answered by investigating the effect of parental background on the two assignment variables for track placement in the Netherlands. These two assignment variables are the same as the ones used in Chapter 4: an elementary school exit test and a track recommendation of the elementary school teacher. The results show that parental background indeed has an effect on both the assignment variables and that it is mostly the high educated parents who are able (consciously or not) to generate better outcomes for their child on the elementary school exit test and the teacher recommendation.

Chapter 6 considers the effects of tracking on a different type of inequality, the inequality caused by month of birth. Just like children of low parental background, since relatively young students perform worse than relatively older students when the track placement decision is made, relatively younger students are more likely to go to the low track (e.g. Jürges and Schneider, 2007; Muhlenweg and Puhani, 2010; Dustmann *et al.*, 2014). Within each class students of different ages are grouped together: The youngest and oldest differ up to 11 months in age. At early ages this difference has been shown to largely explain differences in non-cognitive development (see for an overview Stipek, 2000). The effects on outcomes are caused by absolute age differences at start of schooling and relative age difference within the class. Two potential channels for the effect of relative age are that due to less maturity at school start, the relatively young students develop an early performance lag which might take time to catch up. This early performance lag might also affect the school motivation of the relatively young causing more long term disadvantages. As a consequence of this relative age effect, relatively young students are more likely to be deemed lower

ability students and thus are more likely placed in the low track. At later ages the relative age effect is smaller since the age difference of 11 months is relatively smaller for 15 year olds than for 10 year olds (Crawford *et al.*, 2010). Chapter 6 makes use of cross country variation in the age of selection into tracks and shows that early tracking indeed leads to a stronger relative age effect. However, it also leads to different ability distributions in the low and high track between countries that track early and those that track late. Countries that track late are better able to separate low and high ability students in the low and high track than countries that track early. Therefore, in countries that track early the low track is more heterogeneous in ability. Over time, relatively young students in the low track in countries that track early tend to outperform the relatively older students because at later ages the relative age effect has worn off and their true ability level is revealed. This advantage for the relatively young students is a disadvantage for the relatively old student: They earn less in adulthood and even are more likely to belong to the bottom ten percent earners. Due to the longer lasting relative age effects, countries that have early tracking have higher levels of inequality based on month of birth at later ages.

Chapters 2 to 6 each contribute to answering the two questions this thesis aims to answer: (1) Does tracking increase the average cognitive development? and following from that (2) Does tracking increase inequality among students? Chapter 7 concludes and provides directions for further research. Each of the Chapters 2 to 6 can be read independently of the others and each provides background information on the relevant aspects of tracking students in secondary education. For those new to the topic, the next section provides a basic background on tracking.

1.3 Background on tracking

Tracking is a system wide education characteristic, in the sense that all schools in a system (often comprising a country or state) adhere to it and how tracking is conducted is decided upon at the national or state level. In current day western education systems, tracking only takes place in (post-)secondary and tertiary education (OECD, 2010). No western country tracks students in elementary school. Although all tracked systems formally group students into different educational programs, other differences between tracked systems exist (OECD, 2010). In some countries children are separated into tracks at the age of 10 (Germany and Austria), while in other countries children are tracked later, for instance at age 16 (Sweden and the US).² In a tracked education

2. In the twentieth century a number of countries changed the tracking regimes in their education systems: they either postponed the age of tracking or they lowered the age at which tracking takes place. Examples of such changes are Sweden, and England who moved to tracking at age 16 in the sixties, while in the eighties in Flanders Belgium a policy shift to an earlier tracking age was made, essentially increasing the amount of

system, two or more different tracks are available for students with the number of available tracks being dependent on the system.³ The curriculum differs between tracks, with some tracks offering solely vocational education; while at the other extreme tracks can consist of solely high ability level general education in so called grammar schools or gymnasiums (OECD, 2010). Since the number of tracks differs between countries and also the curriculum content is different among tracks, tracks cannot be easily compared across countries.

In some countries students of all ability levels are kept together in one class, in other words no tracking takes place. Education systems that do not track students are called comprehensive systems. In these systems schools might choose to still group students based on ability using ability grouping or streaming (OECD, 2010). Ability grouping is when students within schools are grouped on ability, for instance by allowing some students to participate in advanced placement classes or university preparatory classes. Streaming is when students are grouped within classes based on ability, for instance, by providing extra challenging assignments for those students who finish their course work early or by letting lower ability students use textbooks from previous years. Some examples of studies which look at ability grouping are Betts and Shkolnik (2000), Figlio and Page (2002), and Duflo *et al.* (2011). Ability grouping and streaming are not considered in the thesis. Comprehensive education systems are studied in this thesis insofar as they act as the counterpart of education systems that do have tracking.

Tracking students in secondary school creates a different environment for students than if students are kept together as they are in a comprehensive system. Three main differences exist between a tracked education system and a comprehensive education system: A tracked education system will have more homogenous classes, different curricula between the tracks, and it might have unequally allocated resources between the tracks. When classes are more homogenous in ability, teachers might find it easier to teach the class. To ensure most students fully comprehend their teacher, teachers logically adapt their teaching to the average student in their class. In teaching practices where a lot of centralized teaching time is required, teaching a homogenous class will allow more students to learn since students are closer to this average ability. In teaching practices where only a limited amount of classroom teaching is required, a homogenous class might still be positive for the teacher since the teacher will have to adapt less to each student that needs an explanation.

differentiation (Braga, Checchi, Meschi, 2013). Changes in the tracking regimes often followed political and societal debates around similar questions as analysed in this thesis.

3. The use of special needs education is often not considered a form of tracking and is therefore also excluded in the analyses in the thesis.

Homogenous classes will ensure that the student's peers are more similar in ability. Besides more homogenous peer groups, if tracking is done on ability, it will also lead to the removal of the very bright peers for the students in the lower tracks and the removal of the not so bright students in the higher tracks. The average peer ability will therefore be higher in the higher tracks than in the lower tracks. Since ability and parental background are highly correlated, tracking on ability also means to some extent tracking on parental background. So besides a more homogenous peer group in ability, tracking will possibly lead to classes being more homogenous on parental background. This peer composition can have effects on student outcomes, and is one of the channels through which the influence of tracking can be seen. Sacerdote (2011) provides an overview of all the different peer effects: From the linear peer effects which averages out across tracks, to the "shining light"/"bad apple" which point to potential beneficial/detrimental effects of very good/very bad peers respectively. Which peer effect dominates is an empirical question not addressed by this thesis. However, it should be noted that also in comprehensive systems homogenous peer groups can form. This can either be due to ability grouping or streaming, caused by religious criteria or by spatial segregation.

Two other differences between a tracked system and a comprehensive system are the differences in curriculum caused by different ability levels and the potential differences in resources. In a tracked education system each track will have a different curriculum, often ranging from more vocationally orientated courses to more general courses. A different curriculum could also have the form of a slower pace at which the same curriculum is given to students. Provided the curriculum is designed with the average ability level of the students of the track in mind, the different curricula should enhance learning in all tracks.

Differences in resources could arise when more (governmental or private) budget is allocated to specific tracks. It could either be that more resources are allocated to the low track since the students in the low track need more assistance, or it could be that more resources are allocated to the high track since it is more prestigious. However a difference in resources is not necessarily solely based on monetary resources, but could also be due to regulations, for instance on teacher credentials. Related to the previous point, if the high track is accompanied by a more demanding curriculum as is often the case, education authorities might require the teachers in that track to have a higher degree than teachers in the other track. This will lead to higher educated teachers in the high track as compared to the low track. It will also raise the costs of teacher salaries in the high track which will lead to a net decrease in resources for other inputs. It remains an empirical question whether more resources are indeed allocated to specific tracks. However, some papers indicate that resources have little to

no effect on educational outcomes (Hanushek, 1986; 1997) in which case this channel of tracking would not be expected to be very important.