7

Conclusion

7.1 Summary of the main findings

This thesis contains studies into multiple aspects of tracking students in secondary education. The first two chapters after the Introduction relate to the effects of tracking on student performance and inequality and how implementation of tracking plays a role in these effects. Chapter 2 shows a positive effect of tracking on student performance at age 15 in European countries when using an instrumental variable approach. Chapter 3 uses a sample of OECD countries and compares countries with no tracking (comprehensive systems) with countries that do track and have up to five tracks available to students. When school principals consider prior performance when accepting the student to the school and students are in a tracked system with four or five tracks, tracking has a positive relation with student's performance. The intuition behind these results is that if prior performance of students is taken into account when deciding on track placement less misallocation of students across tracks occurs and tracks are more homogenous in ability. This seems especially the case when a country allows for enough formal differentiation through tracking by providing students with more than three tracks. The larger number of tracks and track placement based upon prior performance leads to a large number of homogenous tracks, which might very well be the reason for the better outcomes of the students. As discussed in the Introduction, homogenous classes might aid the teacher in maximizing the learning of all students in the class as students are closer to the average ability level the teacher is aiming for. Using prior performance when accepting the student to the school also leads to a lower parental background (PB) effect since the influence of parents is lower.

Chapter 4 provides an answer to the question whether being in the high track is good for outcomes of the marginal student and reports that for some outcomes it is. In the Netherlands the marginal student who is on the threshold of being allowed or not to go to the high track has higher reading and IQ scores and a higher perceived probability to obtain the degree when this student goes to the high track. For other outcomes, as mathematics scores or personality, being in the high or low track has no impact.

Chapter 5 and 6 look at two specific forms of inequality: Inequality due to PB and inequality due to month of birth. Both types of inequalities are well documented in western education systems. However, Chapter 5 looks at whether there is an additional PB effect on the two assignment variables for track placement in the Netherlands over and above the PB effect on ability. Additional influence on track assignment variables, in this case an elementary school exit test and the elementary school teacher track recommendation, will increase the inequality in a system. In the Netherlands especially for students from high educated parents an additional PB effect on the track

assignment variables is found. Chapter 6 does more than repeat the finding that relatively young students are more likely to be placed in the lower track. It adds to the literature the result that this is even more so the case in countries with early tracking, as opposed to late tracking. Due to this increases probability to go to the low track for relatively young students, the ability distribution of the low track becomes wider since more high ability relatively young students are placed in the low track. This provides these students with an advantage since they are the best performers of the track. This has negative effects for the relatively old students in early tracking countries: They have lower earnings at later ages and even are more likely among the bottom ten percent earners.

7.2 Conclusions

From the studies presented in this thesis two overarching conclusions arise. First, tracking students into secondary school is not by definition bad for student outcomes. Previous literature has presented conflicting evidence which pointed either to a positive (e.g. Ariga and Brunello, 2007, Brunello and Checchi, 2007), a negative (e.g. Van Elk *et al.*, 2011; OECD, 2010; Hanushek and Woessmann, 2006) or no (e.g. Jakubowksi, 2009; Pekkarinen, 2008) relation of tracking with student performance and inequality. This thesis adds to this debate, for instance, by looking at specific circumstances in which tracking could be beneficial or not. Chapters 3 to 5 of this thesis suggest that the implementation of tracking is important in achieving desired outcomes and that also the degree of tracking matters. Contrary to much of the policy discussions, tracking can be positive for student performance and equality when track placement is done based on prior performance and a country has more than three tracks. These two elements cause the tracks to be more homogenous in ability, which is potentially the cause of the positive outcomes.

However, the use of prior performance to decide on track placement, does not guarantee optimal allocation of students across tracks. For an optimal allocation of students there needs to be an optimal mapping from the prior performance measure to the different tracks. Chapter 4 looks at this mapping in the Netherlands and finds that the marginal student is to some extent better off in the high track and thus in this specific case an optimal mapping has not been reached. This does not lead to the recommendation that many more students should be allowed to enter the high track, since in Chapter 4 only the effects for the marginal student was looked at. Whether allowing more lower ability students to enter the high track, or reversely, whether allowing more of the higher ability students to leave the low track for the high track has any effects on the peers in each track has not been looked at. Nor has it been looked at what the effects for the marginal student might be if many more marginal

students would also switch from the low to the high track. Chapter 4 does show that at least for students near the threshold being admitted to the higher track causes better outcomes. This suggests that the high track, with its different curriculum, higher ability peers, and perhaps different resources, is also beneficial for at least a specific subset of students in the lower track. One way to increase outcomes for this group is to recreate the same circumstances for them as for those in the high track. As far as policy is concerned, Chapters 3 and 4 highlight that adopting a new policy and implementing it are two different things.

Inequality based on parental background is very common and in Chapter 3 it is argued that this inequality is lower in education systems that track students and base track placement on ability as measured by prior performance. In these systems parents are only able to influence the track placement of their children by influencing the performance of their children and not by influencing the outcomes (track placement) directly. However, Chapter 5 provides evidence that considering prior performance to decide on track placement does not fully remove the effects of parental background. In the Netherlands, where at the time of this research an elementary school exit test and an elementary school teacher track recommendation were mandatory to decide on track placement, an additional effect of parental background is found on the two track placement assignment variables. This shows that when parents are indeed limited in their direct influence on the outcomes of their children, they might find ways to increase their indirect influence. This might be a conscious choice, or could be considered regular parenting: Parents will almost always seek out the best outcomes for their children and will, if direct influence is not possible, prepare them to obtain these. Thus requiring the use of prior performance to decide on track placement is no silver bullet that ensures improved performance and lower inequality. Just like with the considerations of using prior performance to decide on track placement as outlined in the paragraph above, using prior performance will also not likely remove the full effect of parental background.

The second conclusion, which arises from Chapter 6, is that relative age should be taken into account when formulating the mapping from prior performance to tracks. Relative age, or month of birth combined with the national cutoff date to start elementary school, influences track placement in that it introduces a bias in the prior performance of relatively young students. As a consequence from this bias the relatively young students are more likely to be placed in the low track. In turn, due to distributional shifts, this leads to a disadvantage for the relatively old in educational outcomes at age fifteen and in the labor market. To prevent this from happening, multiple things can be done. First, the relative age bias can be ruled out, for instance, by using different performance thresholds for relatively young and old students or by

using an age correction on those performance measures used for track placement. Second, school admission can be done by a rolling admission, although this might be difficult to accomplish. Third, stimulation of track mobility when the first track allocation turns out to be incorrect might repair the earlier damage done by misallocation. Fourthly, the age at which tracking takes places can be postponed until the relative age effect no longer affects track placement. It must be said however that our results show that also in education systems that track late, relative age has an effect on later outcomes. Thus both early and late tracking countries are confronted with inequalities caused by the birth month of students.

7.3 Limitations

To ensure robust findings, in the chapters of this thesis numerous checks were conducted. Still, caution must be taken in interpreting the findings from this thesis. All chapters come with their own dis- and advantages and all are dependent on the specific circumstances and data availability. Chapter 2 uses a sample of European countries and presents average effects from these countries. Furthermore, it uses an instrumental variable approach which, in the case of bad instruments, could in serious cases lead to more biased estimates than OLS. The exclusion restriction, which aims at preventing this bias, is tested and indeed more aspects than the level of tracking in education systems are affected by the instrument. This should be kept in mind when interpreting the results. However, the direction and size of the effect estimated with IV does not deviate that much from the OLS estimates and are also to some extent in line with previous studies. Chapter 3 is based on a sample of OECD countries. To avoid relying solely on cross country variation, the main models are supplemented with within country models. This is done to alleviate the country heterogeneity from cross country analyses. When students with different ability levels would sort between schools where the school principals do or do not consider prior performance on accepting the students to the school, this could influence the results. Therefore also models using solely the variation of the use of prior performance by school principals between countries are presented. The results from Chapter 4 and 5 are specific to the Dutch context as data from the Netherlands used and must be interpreted as such, although the results could be indicative for other contexts. Chapter 6 uses data from OECD countries and therefore presents average effects and is not able to zoom into country specific circumstances.

7.4 Future research

This thesis contributes to a deeper understanding of the effects of tracking on student performance and inequality. However, it also opens up new questions. The analyses in

this thesis raise questions of great consequence for policy makers and other stakeholders: What is the optimal number of tracks? What are the optimal thresholds for prior performance to enter into these tracks? What is the optimal age of tracking to ensure the relative age bias is not affecting track placement? Or how should prior performance be corrected for to eliminate the relative age bias in track assignment measures? Also questions on the effect of curriculum differences or curriculum changes can ensure a deeper understanding of the different effects that tracking secondary school students might have.

The chapters of this thesis have not discussed all questions related to tracking, student performance and inequality. This thesis focused almost exclusively on cognitive outcomes and inequality, and to some extent on non-cognitive and labor market outcomes. Other outcomes, like civic outcomes, were not taken into account but might very well be important for the stakeholders in education. And tracking students in secondary education might have an impact on the development of those outcomes.

Furthermore, since currently the only cross nationally cognitive tests for students are on general education subjects (languages, mathematics and science), it might be that comprehensive systems have an advantage in cross national research as compared to tracked systems. In tracked education systems, especially those with vocationally orientated tracks, part of the student population might encounter little general education in secondary school, leading automatically to worse cognitive outcomes on these subjects. However, the vocationally oriented students might very well outperform students in other countries in vocationally oriented tests. Whether this is the case and how this influences the debate on the performance differences between comprehensive and tracked education system is open for further research.

As touched upon in the Introduction, little attention is given to the different channels through which the effects of tracking reported in this thesis run: The homogenous classes which are easier to teach but have different peers, the curriculum effects and possible resource differences. Future research should also shed more light on these topics which were not investigated in the chapters of this thesis.

Finally, the identification strategies used in this thesis are the best available given the data and the questions. This does not imply that the answers given to the questions in this thesis are final. Like in any scientific effort, they will have to be replicated using data from other regions and using different methodologies.