Chapter 8

Conclusions and discussion

In this dissertation I have investigated the changing association between family background and educational attainment in the Netherlands during the 20th century. This association is a measure of the inequality in access to education, as it indicates the extent to which persons with a more privileged background are more likely to attain a higher level of education than persons with a less privileged background. This inequality in access to education is not only important to investigate because education is a valuable and scarce resource in its own right, but also because it influences future success in other domains of life, like work, family formation, and health. The research literature on the inequality of access to education has a long history (Hout and DiPrete, 2006; Breen and Jonsson, 2007). This dissertation contributed to this literature by studying the following aspects of inequality in access to education: 1) the inequality in the outcome of the process of attaining education, 2) the inequality during the process of attaining education, as well as the relationship between these two types of inequalities. I have labelled these two types of inequality Inequality of Educational Outcome (IEOut) and Inequality of Educational Opportunity (IEOpp) respectively. The overarching research question that guided the individual studies that make up this dissertation has been: "To what extent, how, and when has a trend toward less inequality in educational opportunities and in educational outcomes of persons from different family backgrounds occurred in the Netherlands?"

As a point of departure I replicated in Chapter 2 a study by De Graaf and Ganzeboom (1993) using more, and more recent data. This replication served as a benchmark, as it represents what can be learned from the most recent data using 'default' methods. The remaining chapters consisted of applying new methods that improved on these 'default' methods. Chapters 3, 4, and 5, showed three ways of improving the estimates of IEOut: In Chapter 3 a scale of education was empirically estimated to replace the *a priori* scale that has been used in the 'default' method. In Chapter 4 the trend in IEOut was estimated using a local polynomial curve which is more flexible than the quadratic curve and more powerful than the discrete curve that have been used in the 'default' approach. In Chapter 5 a new method was introduced for testing whether the relative differences in effects of occupational status and education of the father and the mother on the offspring's educational attainment have changed over time. Chapter 6 showed a new way of relating IEOpps to IEOut, which also turned out to provide a meaningful way of analyzing the effect of educational expansion on IEOut. Chapter 7 showed a way of improving the estimates of the IEOpps, by proposing a sensitivity analysis to assess the potential impact of unobserved variables on the results.

The conclusions from all these chapters will first be discussed in detail, and are then summarized by answering the overarching research question. Finally, some shortcomings of these studies are discussed together with some recommendations for future research.

8.1 Conclusions

8.1.1 A replication

The dissertation started with a replication of the study by De Graaf and Ganzeboom (1993), which was the Dutch contribution to an influential international comparative project by Shavit and Blossfeld (1993). The role of the replication in this dissertation is to create a point of reference in terms of the estimated trend in inequality of access to education using 'default' methods. De Graaf and Ganzeboom (1993) studied IEOpp and IEOut, which both play a prominent role in this dissertation. Moreover, the data used in this dissertation is an extension of the data used by De Graaf and Ganzeboom (1993). They used data from ten cross-sectional surveys that were post-harmonized and then stacked to form a single dataset. Ganzeboom and Treiman (2009) have since extended this data as part of the International Stratification and Mobility File (ISMF) such that the Dutch part of this file now contains information from 54 surveys. It is this data that has been used throughout this dissertation.

The main finding of this replication is that despite the fact that this replication used more than five times as many respondents (69,868 versus 11,244 respondents) and covered 20 additional years (1891–1980 versus 1891–1960), the results remained largely unchanged. Using default methods on the extended dataset the following trends in IEOpp and IEOut were found for the Netherlands: a significant negative trend in IEOut and a significant negative trend in IEOpp for the transition whether or not to continue after primary education; mixed evidence for the trend in IEOpp for the choice of track during secondary education; and no trend and in some cases a positive trend for the transition whether or not to finish tertiary education. Moreover, these trends are mostly found to be linear.

8.1.2 IEOut: operationalizing education, the trend, and family background

Chapters 3, 4, and 5 of this dissertation proposed three ways of improving on the ‘default method’ of estimating the trend in IEOut.

Chapter 3 focussed on the values assigned to each of the educational categories. These values are necessary in order to estimate IEOut. Following De Graaf and Ganzeboom (1993), the replication in Chapter 2 assigned values to educational categories by distinguishing between four educational categories (primary, lower secondary, higher secondary, and tertiary education) and assigning them the values 1 to 4. A major advantage of this method is that it is easy to apply, all that is necessary is a rank order of the educational categories. A disadvantage is that this method implicitly assumes that the distances between successive educational categories are all equal. An often-used alternative approach is to assign each category a value equal to the number of years it would take a ‘standard’ student to complete that category. An advantage of this method is that these values can easily be derived for most educational systems from (semi-)official documents. However, it conflates two distinct concepts: the duration and the value. As a result, the rank ordering of educational categories based on these standard durations sometimes does not correspond to the rank ordering based on a priori knowledge about the values. This is the case in the Netherlands for higher secondary vocational education [MBO], which has led Ganzeboom and Treiman (2009) to apply an *ad hoc* adjustment to their scale values when creating their scale for the ISMF. Another potential problem with these *a priori* scales of education is that they assume that the values of the educational categories have remained constant over time, while there are two plausible mechanisms through which the value of an educational category could change over time: educational reform, which can mean that an educational category before and after a reform should be treated as two different categories, and changes in the supply of highly schooled labor relative to the demand for highly schooled labor, so called ‘diploma inflation’.

Chapter 3 improved these standard ways of assigning values to the educational categories by empirically estimating a scale of education. This scale of education is estimated such that it is optimal for predicting occupational status, using a model with parametricaly weighted covariates proposed by Yamaguchi (2002). The resulting scale largely corresponds with the *a priori* scale by Ganzeboom and Treiman (2009). The major deviation from the *a priori* scale is that the *a priori* scale overrates the value of lower secondary vocational education [LBO], which means that respondents with LBO had, on average, lower status occupations than was predicted using the *a priori* scale. The resulting scale also showed that there is little evidence that the values changed over time, as measured by the year in which the survey was held. The time at which the survey was held was used as a proxy for the time when the respondents held their job. As a consequence, the lack of change over time is an indication that changes in the labor market during the period that was studied (1958 to 2006) had little effect on the relative distances between educational categories. However, the values of two educational categories did differ when comparing cohorts that were in education before and after a major educational reform in the Netherlands, the "Mammoet Wet" or "Mammoth Law", which was implemented in 1968. The categories that were influenced by it were lower general secondary education [MAVO] and higher professional education [HBO]. The change in the value of MAVO was to be expected, as this diploma changed from a level that prepared for the labor market to a level that prepared for a subsequent level of education (MBO). A possible reason for the change in the value in HBO could be that it became accessible from higher general secondary education [HAVO].

Showing the consequences of using this new scale rather than the *a priori* scale for the estimates of IEOut was one of the subsidiary aims of Chapter 4. The main aim of this chapter was to assess whether or not the trend in IEOut has changed over time. Past research had found a steady negative trend in IEOut, and found no evidence for any non-linearity in this trend. However, it is implausible that this linear trend will continue as this would eventually result in a negative association between family background and educational attainment. So, at one point in time the negative trend in IEOut will have to slow down, and the aim of Chapter 4 was to try to detect this declaration of the trend. This chapter hypothesized that the lack of evidence for a non-linear trend in the default approach was due to the methods used in testing for non-linearities: these methods either estimated a non-linear trend using a quadratic trend, which could be not flexible enough to adequately detect any non-linearities in the trend, or as a discrete trend model, which could be too flexible and thus not powerful enough. The alternative proposed in this chapter was to represent the trend as a local polynomial curve, which is more flexible than a quadratic curve but more powerful than a discrete curve.

This chapter did find evidence that the trend has been non-linear, but did not find the expected deceleration in the decreasing trend in IEOut. A period of negative trend was found for both men (1941–1960) and women (1952–1977), which was preceded by a period of significantly accelerating trend (1935–1944 for men and 1949–1952 for women). There is some evidence — only for men — that the negative trend decelerated prior to becoming insignificant, but this deceleration is not (yet) significant. There is no indication that the negative trend for women decelerated prior to becoming insignificant, indicating that the lack of significance of the negative trend in the youngest cohorts has more to do with lack of power than with a lack of negative trend. Surprisingly, the trend did not show any effect of a major educational reform, the

'Mammoet Wet' or 'Mammoth Law', which was aimed at reducing IEOut and was implemented in 1968.

A subsidiary aim of this chapter was to assess the robustness of these conclusions to three potential sources of error: different scales of education, differences in quality of the data across surveys, and missing data. Using the scale of education estimated in chapter 3 rather than the *a priori* scale by Ganzeboom and Treiman (2009) led to a slightly more stable trend, but did not qualitatively change the conclusions. Controlling for differences between surveys led to a decrease in trend for the earliest cohorts, while using multiple imputation to control for missing values did not influence the results.

In Chapter 5 I assessed which resource — occupational status or education — and which parent — the father or the mother, the highest educated/status parent or the lowest educated/status parent, or the parent with the same sex as the respondent or the parent with a different sex to the respondent — contributed most to the offspring's educational attainment. The results indicate that the distinction between highest and lowest status parent is the main distinction between the parents, rather than the distinction between fathers and mothers or the distinction between the parent with the same sex as the respondent or a different sex to the respondent. There is also moderate evidence that occupational status is more important than parental education. I also found that the mother being a homemaker had a negative effect on the educational attainment of the offspring if the mother has little education and the father has a low status job, but that this effect becomes positive when the mother is well-educated or when the father has a high status job.

In this chapter I also investigated whether the relative contributions of each of these resources changed over time. I expected the value of the contributions of the mother's resources to have increased over time relative to the values of the resources contributed by the father due to changes in the roles of men and women in society during the period studied (1939 till 1991). I also expected the value of occupational status to decline as it is more closely related to the economic resources available in the family, and economic constraints have become less likely to limit the educational choices as almost everybody has become wealthier and education has become more heavily subsidised during the period studied. In order to test these hypotheses, I used a model with parametricaly weighted covariates proposed by Yamaguchi (2002), which estimates the model under the null hypothesis that the relative contributions of these resources have remained unchanged over time. Contrary to my expectations, this hypothesis could not be rejected.

8.1.3 Combining IEOpp and IEOut

When investigating inequality in access to education, it is useful to distinguish between inequality during the process of attaining education (the IEOpp) and the inequality in the final outcome of that process (the IEOut). It is also useful to recognize that IEOpp and IEOut provide complementary information; a discussion of the process of attaining education can be meaningfully supplemented by a discussion of the outcome of that process and *vice versa*. In order to make the best use of this complementarity between IEOpp and IEOut, one needs to move beyond separate discussions of IEOpp and IEOut and towards an integrated discussion of the two. Chapter 6 proposed a new method that makes such an integrated discussion possible. This method starts with the standard model for estimating IEOpps, the sequential logit model as proposed by Mare (1981), which estimates the effect of family background on the probabilities of passing from one level of education to the next. It then shows that this model implies a decomposition of IEOut as a weighted sum of the IEOpps, such that the weights assigned to each transition between levels of education are the product of three elements: 1) the proportion of respondents at risk of passing that transition, which means that a transition receives more weight when more people are affected by it; 2) the variance of the indicator variable indicating whether or not respondents passed that transition, which means that less weight is given to transition where virtually everybody fails or virtually everybody passes; and 3) the expected increase in highest achieved level of education due to passing that transition, which means that a transition receives more weight when passing it is more profitable. This makes it possible to supplement the IEOpps with estimates of how relevant these IEOpps are for IEOut. Moreover, it provides a substantively interpretable mechanism through which educational expansion can influence educational inequality, as educational expansion influences the probabilities of passing the educational transitions, which influence the weights, which in turn lead to changes in IEOut.

When applying this methodology to the Netherlands, I distinguished four transitions: continue or not after primary education, taking the vocational track versus the academic track, continue to higher secondary vocational education given that a respondent entered the vocational track, continue to university given that a respondent entered the academic track. I found that the latter two transitions not only have low IEOpps, which was already known, but they also have low weights. These low weights were primarily due to the relatively low proportion of respondents that are at risk of passing these higher transitions compared to the lower transitions. By contrast, educational expansion had a big influence on the first two transitions. The first transition, whether or not to continue after primary education, started out as the main source of IEOut, but declined quickly as passing this transition became almost universal. The second transition, whether to enter the vocational track [LBO and MAVO] or academic track [HAVO and VWO], strongly increased in importance as the percentage of people passing that transition increased to about 50%, which resulted in an increase in the variance of the dependent variable, and as more and more people became at risk of passing this transition.

8.1.4 IEOpp: the influence of unobserved variables

The standard model for estimating IEOpps, the sequential logit model, has been subject to an influential critique by Cameron and Heckman (1998). They argue that, like any other model, a sequential logit model cannot include all variables that influence the dependent variable. However, leaving these variables out will influence the results, even if these variables are not confounding variables. These so-called unobserved variables influence the results through two mechanisms. First, the 'averaging mechanism' is based on the fact that when a variable is left out of the model, one models the probability of passing the transitions averaged over the variables that are left out. As a consequence, leaving the unobserved variable out of the model will lead to estimates of effects of the observed variables on the *average* probability of passing within groups defined by the observed variables rather than the effects on the *individual's* probability of passing. These two are different in non-linear models like logistic regression because the unobserved variables are related to the probabilities through a non-linear function. Second, the 'selection mechanism' is based on the fact that a variable that is not a confounding variable at the first transition is likely to become a confounding variable at later transitions. The reason for this is that the process of selection at the earlier transitions will introduce correlation between the observed and unobserved variables at the later transitions.

This suggests that one needs to control for these unobserved variables, but it is by definition impossible to get an empirical estimate that is controlled for variables that have not been observed. However, it is possible to create a scenario, by specifying assumptions about the unobserved variables, and estimating the effects within that scenario. There are roughly two ways in which these scenarios can be used. First, one can try to put as much empirical information as possible into these scenarios. For example Mare (1993, 1994) uses the similarity between siblings to capture the unobserved variables on the family level. Alternatively, one can use a set of scenarios to assess the sensitivity of the estimates to the assumptions. Chapter 7 is an example of this latter approach as it proposed a set of scenarios that is useful for such a sensitivity analysis and a method for estimating the effects within these scenarios. This method was illustrated by replicating the analysis in Chapter 2, showing that the results of statistical tests were robust to changes in the assumptions about unobserved heterogeneity, but that the effects of both father's occupational status and father's education were likely to be underestimated, as these effects were stronger in scenarios with more unobserved heterogeneity. Scenarios with more unobserved heterogeneity also resulted in a stronger downward trend over time in the effect of father's occupational status and education, indicating that the trend in the effects of parental background variables across cohorts is also likely to be underestimated. However, the effect of father's occupational status and education decrease less over transitions in scenarios with more unobserved heterogeneity. This indicates that the commonly found pattern of decreasing effects of family background variables over transitions is at least in part due to unobserved heterogeneity.

8.1.5 Summary

These conclusions can be summarized by explicitly answering the overarching research question: "To what extent, how, and when has a trend toward less inequality in educational opportunities and in educational outcomes between persons from different family backgrounds occurred in the Netherlands?" The answers to this question can be broken up into the following elements:

IEOut

- The trend in IEOut was shown to have decreased during the third quarter of the 20th century, during which time it approximately halved. This negative trend was preceded by an acceleration of the trend, and there is some indication that IEOut was initially increasing. The non-linearity of this trend is a new finding, as previous studies failed to reject the hypothesis of a linear declining trend.
- The 'Mammoet Wet', a major educational reform in the Netherlands implemented in 1968, did not have a noticeable influence on IEOut.
- An improved scale for the educational categories was created in this dissertation, but this was found to have little effect on the estimated trend in IEOut.
- The relative contributions of the education and occupational status of the father and the mother to the respondent's educational attainment were found to have remained constant over cohorts.

IEOut and IEOpp

- The main driving force behind the trend in IEOut turned out to be the major shift in which transition between educational levels contributed most to IEOut. Initially, the transition between whether or not to continue in education after finishing primary education was the main contributor to IEOut. However, the contribution of this transition quickly declined as passing this transitions became almost universal. At the same time the contribution of the second transition between entering the academic or vocational track increased in importance as more people became at risk of passing that transition and as the number of people entering the academic track and vocational track became more evenly balanced. This shift between the transitions resulted in both the initial increase in IEOut, as the decline of the contribution from the first transition was more than compensated by the increase of the contribution from the second transition, and the subsequent decline in IEOut, as the less unequal second transition replaced the more unequal first transition as the dominant source of inequality.

IEOpp

- At the lowest transitions a declining linear trend in the IEOpps over time was found, while at the higher transitions the evidence became mixed with negative, insignificant, and even positive trends.
- The IEOpps at the lower transitions were higher than the IEOpps at the higher transitions.
- A sensitivity analysis showed that qualitative conclusions are robust, but that both the size of the IEOpps and the size of the trend are likely to be underestimated when the unobserved variables are not accounted for.

8.2 Discussion

What all chapters in this dissertation have in common is that they used data from the International Stratification and Mobility File [ISMF] (Ganzeboom and Treiman, 2009). As a consequence, all these chapters share the strengths and weaknesses associated with this source of data. One of these weaknesses is that the ISMF contains data from surveys of differing quality. Chapter 4 found that controlling for differences between surveys did have a moderate effect on the estimated trend in IEOut. Future research could extend on this finding by also modelling the effects of survey characteristics, thus gaining more insight into the way survey quality influences the substantive conclusions that can be drawn from it. This would turn the variation between the surveys present in the ISMF from a potential weakness into a strength, as this variation can then be used to control for characteristics of the survey in ways that are impossible when analyzing surveys separately or only analyzing surveys with certain (high quality) characteristics.

Another potential weakness is the way time is measured using so-called synthetic cohorts, that is, cohorts that are observed in a cross-sectional survey. These synthetic cohorts are used to estimate the trend in IEOpp and IEOut, and thus play a key role in this dissertation. The key advantage of using synthetic cohorts is that it makes it possible to study a long period of time using a large amount of data. However, there are also problems associated with the use of synthetic cohorts. The first problem is that a synthetic cohort is not a proper sample from the population of people born in a certain year, but a sample from the population of people born in a certain year *and* who are still alive and living in the Netherlands at the time the survey was held. This can be a problem for cohorts that are very old when the survey was held because in these cohorts higher-educated respondents are likely to be over-represented, as higher-educated persons are more likely to live longer. Such a selection on the dependent variable can bias the results (Breen, 1996). This was partly solved in most chapters by only using respondents younger than 65 years old[1]. This way, not enough people will have died for this to have become a problem. The second problem with synthetic cohorts is that education happens over a period of time, so it is not exactly clear which historical period is represented by a cohort. A reasonable choice is to look at the time when the respondent was 12, as in the Netherlands that is the age at which people make the most important decision in their educational career, but any such choice will necessarily be an approximation. This is particularly relevant when studying the consequences of a policy change, as synthetic cohorts will only approximately classify the respondents as being affected or not affected by the policy change.

Another difficulty with the use of cross-sectional surveys like the ISMF is that they do not directly measure which transitions a respondent passed. The transitions a respondent has made are reconstructed based on the respondent's highest achieved level of education and a simplified model of the educational system. In particular, in order to be able to reconstruct a respondent's educational career, such a model must impose that a respondent can only reach a certain level of education through one route. This is a limitation, especially within educational systems consisting of multiple tracks, as it precludes the study of indirect paths through the educational system. This can be of substantive interest as these indirect paths represent 'second chances' open to respondents after they have chosen/been placed in a certain track. As a consequence, the 'synthetic educational careers' in the ISMF preclude the study of the question concerning who benefits most from these second chances: the people

[1] Exceptions are chapters 2 and 7, which replicate the study by De Graaf and Ganzeboom (1993).

with fewer family resources, who were initially disproportionately placed in the lower tracks, or the people with more family resources, who are better capable of making the best use of any loophole in the system.

A way to avoid problems with synthetic cohorts and synthetic educational careers is to use data for which the time at which events took place and the educational career are directly measured. This type of information is available in panel data, where students are followed during their educational career, or in cross-sectional surveys where respondents are asked to retrospectively reconstruct their educational career. However, this does not mean that these sources of data are uniformly better than cross-sectional surveys that only asked for the highest achieved level of education. It is actually striking how much the strengths and weaknesses of these different types of data complement one another. An analysis of panel data and retrospective career data can add to an analysis of highest achieved level data as the panel data and retrospective career data have directly observed time and educational careers. An analysis of the final stages of the educational process and the outcome of the educational process is difficult to make in the panel studies due to attrition, but neither the retrospective career data nor the highest achieved level data suffer from this problem. The available panel studies contain data on only a few cohorts, making it difficult to get a detailed description of changes over time, while both the retrospective career data and the highest achieved level data contain information on many cohorts. However, the retrospective career data contain data on relatively few respondents, meaning that each cohort consists of a small number of respondents. The panel data contain few cohorts, but each cohort contains many respondents. The highest achieved level data tend to contain many cohorts, and each cohort consists of many respondents. The retrospective career data can suffer from the fact that its information is based on what a respondent can remember of events that, for some cohorts, occurred many years previously. The panel data do not suffer from this as the data is collected shortly after the events occur, while the highest achieved level data collects information on the highest achieved level of education, which is much more salient and easier to remember than the entire educational career. Future research could make real progress if it were to exploit these complementarities between the data sources rather than continuing to use them separately.

On a more general level, a discussion of this dissertation needs to confront its rather specific nature, as one of the defining characteristics of this dissertation is the central role that methodological innovations play in every chapter. One could ask whether such a methodological orientation is a good thing. In the end, methodology is just a tool and not an aim in itself. I think that such a methodological dissertation has its place within a substantive field like social stratification research, but such a study should meet a number of challenges. When proposing new methodologies, it is easy to get carried away and to purely focus on applying the latest and most fashionable techniques. Similarly, when pointing out a defect in a methodology it is very easy to forget that all models are defective, as models by their very nature are simplifications of reality and a simplification is nothing other than a 'reasonable error'. In other words it is not enough to show that one can invent or apply new methodologies or show that some 'old' methodology is defective, one must also show that this helps to either better answer existing questions or answer new questions. Moreover, when one proposes new methodologies it is easy to forget that the aim is to create a new tool that can be used by others. If it takes more than a reasonable amount of effort for other researchers to use this new method, then the methodological study has not achieved its aim. In this dissertation I have attempted to meet these challenges by focussing in each chapter on using the methodological innovations to answer substantive questions, leading to some truly new findings, thus showing that it is not just new technology but that this new technology contributes to the study of educational inequality. Moreover, the methods proposed in this dissertation used either existing software or new software was written to implement the new methodologies. In particular, chapter 4 used the `locfit` module by Loader (1999), while two new software modules were written within the statistical programme Stata (StataCorp, 2007) to implement the remaining new methodologies: `seqlogit` (Technical materials II) for chapters 6 and 7 and `propcnsreg` (Technical materials I) for chapters 3 and 5. This has enabled the new methods proposed in this dissertation to be accessible to other users.