Education, Stratification & Simulation

Rense Corten
Education, Stratification & Simulation

Doctoraalscriptie Sociologie, Universiteit Utrecht

door
Rense Corten
9933557

1e Begeleider: Ineke Maas (UU)
2e Begeleider: Chris Snijders (TUE)

Utrecht, Juni 2004
Voor mijn moeder († 2001)
Dankwoord

## Contents

1. **Introduction** .......................................................................................................................... 1  
   1.1 Status attainment research.................................................................................................. 1  
   1.2 Theories on education & status inequality ................................................................. 2  
   1.3 Stratification and simulation .......................................................................................... 4  
   1.4 Overview ....................................................................................................................... 5  

2. **Theories & Research** .............................................................................................................. 6  
   2.1 Boudon’s theory of social stratification ......................................................................... 6  
   2.2 Expansion, transitions and selection: Mare ..................................................................... 14  
   2.3 Empirical efforts ............................................................................................................. 16  
   2.4 Conclusion ..................................................................................................................... 19  

3. **Research Questions & Simulation Design** ........................................................................ 21  
   3.1 Research questions ......................................................................................................... 21  
       *Inequality of Educational Opportunity* ........................................................................ 22  
       *Selection and IEO* ....................................................................................................... 23  
       *Inequality of Status Opportunity* .............................................................................. 24  
       *The effect of education* .............................................................................................. 25  
   3.2 Simulation design .......................................................................................................... 26  

4. **Simulation Results** ............................................................................................................. 30  
   4.1 Dependent variables ....................................................................................................... 30  
   4.2 Independent variables ................................................................................................... 31  
   4.3 Results: IEO .................................................................................................................. 33  
       *Educational expansion and selection* ....................................................................... 36  
   4.4 Results: ISO ................................................................................................................ 39  
       *The returns of education* ........................................................................................ 41  
   4.5 Conclusions .................................................................................................................. 43  

5. **Testing the Model** ............................................................................................................. 44  
   5.1 Testing IEO predictions ................................................................................................. 44  
       *Data* ............................................................................................................................ 44  
       *Method* ....................................................................................................................... 45  
       *Variables* .................................................................................................................... 45  
       *Results: primary and secondary effects; expansion* ................................................. 48  
       *Results: expansion & selection* .............................................................................. 50  
   5.2 Testing ISO predictions ................................................................................................. 52  
       *Data* ............................................................................................................................ 52  
       *Variables* .................................................................................................................... 52  
       *Results: ISO* ............................................................................................................. 55  
       *Results: Returns of education* ................................................................................. 59  
       *A different way of testing: estimation of effects* ....................................................... 60  
   5.3 Summary ....................................................................................................................... 62  

6. **Conclusions & Discussion** .............................................................................................. 63  
   6.1 Education, opportunity and social inequality ............................................................... 63  
   6.2 The merits of simulation ............................................................................................... 69  
   6.3 Suggestion for further research .................................................................................. 71  

**Literature** .............................................................................................................................. 72
List of Tables and Figures

Table 3.1 Description of parameters used in the simulation ........................................ 29
Table 4.1 Descriptive statistics of dependent variables in the simulation ........................ 31
Table 4.2 Descriptive statistics of independent variables in the simulation, with the values used by Boudon ................................................................................................. 33
Table 4.3 Regression models with IEO as dependent variable (simulated data) ................. 34
Table 4.4 Regression models with IEO as the dependent variable (simulated data) .......... 37
Table 4.5 Regression model with the conditional primary effect at the highest educational level as dependent variable (simulated data) .................................................. 38
Table 4.6 Regression model with ISO as the dependent variable (simulated data) ............ 39
Table 4.7 Regression model with the returns of education as the dependent variable (simulated data) ........................................................................................................ 41
Table 5.1 Dependent and independent variables in four Dutch education cohorts, 1977-1993 .................................................................................................................................. 48
Table 5.2 Observed and predicted IEO, The Netherlands 1977 – 1993 .......................... 49
Table 5.3 Observed and predicted conditional primary effect, The Netherlands 1977 – 1993 .......... 51
Table 5.4 Cohort level dependent and independent variables in the Netherlands, 1946 – 1980 ......................................................................................................................................... 54
Table 5.5 Descriptive statistics of dependent and independent variables in the simulation analysis of ISO ................................................................................................. 54
Table 5.6 Descriptive statistics on dependent and independent variables in five countries, 1946-1980 .................................................................................................................. 61
Table 5.7 regression models with ISO as the dependent variable in five countries, 1946-1980 .......................................................... ...................................................................................... 62
Table 6.1 Hypotheses generated by the simulation model .................................................. 65

Figure 4.1 Primary and secondary effects ................................................................. 33
Figure 5.1 Observed versus Predicted IEO, The Netherlands 1977-1993 ...................... 49
Figure 5.2 Observed and predicted conditional primary effect, The Netherlands 1977 – 1993 ............................................................ ...................................................................................... 51
Figure 5.3 Proportion of high status positions (scoring higher than 63 on the ISEI scale), The Netherlands 1946 -1980 ......................................................................................... 57
Figure 5.4 Level of meritocracy, calculated as unemployment rate + .5, The Netherlands 1946 -1980 ................................................................. ...................................................................................... 57
Figure 5.5 IEO, The Netherlands 1945 – 1980 ................................................................ 58
Figure 5.6 Observed and predicted ISO using different scales for meritocracy, The Netherlands 1946 -1980 ......................................................................................... 58
Figure 5.7 Observed and predicted returns of education, using different scales of meritocracy, The Netherlands 1946 -1980 ................................................................. 60
1. Introduction

Social inequality and stratification have among the main focuses of the discipline of sociology, and still make up one of the largest research fields in the field. Who gets what, which groups manage to get a disproportional large piece of the pie, and why? Inequality is one of the most important organizational principles of society, and the source of many social phenomena. The “pie” can consist of many things: material wealth, status, power, education or health. In this study, it is ultimately social (occupational) status. In their lifetime, people may move from one status level to another; the starting point (social origin) is the occupational status of their parents, the destination the status of their own job as adults. Almost by definition (because status is defined in terms of what is valued most), people strive to get a status as high as possible.

The mechanisms by which status is transferred from one generation to the other are subject to change over time; while status in earlier times used to be mostly hereditary, today’s modern societies appear to be more or less meritocratic: status is rewarded for what people have achieved, rather then for who they are (Young, 1958). In the case of occupational status attainment, the individual characteristic that is generally supposed to be decisive is formal education: your educational credentials determine the kind of job you are going to get. On the other hand, there is supposed to be an important relation between social origin and educational attainment: your social class determines (to some extent) your schooling career. Thus education is regarded as an important intervening variable between social origin and occupational status. The exact nature of this relation however is not really clear yet, and therefore the role of education in the process of status mobility is the main focus of this study.

1.1 Status attainment research

This question is of course not new: within the discipline of sociology there is a large body of research generally indicated as status attainment research. This type of research has a long tradition, probably starting with Blau and Duncan’s study “The American Occupational Structure” (1967). In this classic study they first introduced the now standard five-variable path model to describe patterns in the relation between social origin, education and occupational status. Using this method they were able to show that although the level of education had become an important determinant of occupational status in modern America, social background still had it’s influence as well. This approach has been widely followed in many different contexts: different scholars have expanded the analyses to women, included different countries, used regression analyses instead of correlation coefficients, and looked at the influence of different political regimes on status attainment (see Rijken, 1999). The large majority of the studies in this field have however remained highly empirical, and have, like Blau and Duncan did, avoided the systematical development of theory: status attainment is usually conceived as a relation between individual characteristics that can be empirically estimated. This led Coleman to remark on this type of research:
The body of research in sociology is in “social stratification” or “status attainment”, and it almost completely fails to make the micro-to-macro transition. Ordinarily research in social stratification treats a change of job as if it were an individual decision: The determinants are background characteristics of the individual, aspects of life history that affect occupational mobility. The destination occupation is regarded as unlimited in number of open jobs; taking a new job of a particular type is analysed in exactly the same manner as a change of attitude. Jobs are scarce commodities, however, and a new job is obtained only in competition with others. (Coleman 1987, pp 163.)

The main goal of this study is examine whether a more theoretical approach to status attainment can be found and maintained. The main question therefore is a rather basic one:

What are the mechanisms that generate educational inequality and status inequality?

1.2 Theories on education & status inequality
It is of course an exaggeration to claim that there have been no theories whatsoever in the field of status attainment, and this section will describe some of the theories that have been developed. The theories that will be discussed here are those by Boudon, Thurow and Mare. These are probably the most discussed and most important theoretical contributions in the field although efforts have been made by others, for example by Sørensen (1976).

Boudon’s study Education, opportunity and social inequality was published (in English) in 1974, and immediately caused considerable controversy. The main reasons for this controversy were probably Boudon’s high pretentions, his highly unorthodox methods and his far-reaching conclusions. Boudon set out to provide an explanation for some findings in sociological research that he found to be highly paradoxical, the most famous of these the Lipset-Bendix paradox that states that although educational systems are very different between countries and have changed considerably within countries, there have hardly been any differences between and within these countries in intergenerational mobility. In order to face this and other puzzles Boudon designed a thought-experiment of an ideal-type society, formulated as a numerical example. Using simple assumptions a two-stage model was constructed, consisting of two parts: a hypothetical educational system leading from social origin to educational attainment, and a system of job distribution that leads from educational attainment to social status. In the educational model, the educational system is conceived of as a sequence of transitions, in which a number of students have to take decisions on either leaving or staying in the system at each transition, the outcomes of these decisions being (according to implicit rational choice assumptions) partly dependent on social background. The labour market model then proceeds with assigning the
individuals who have reached a certain educational level to a status position, assuming mechanisms of job competition where the most highly educated are first in line to get to the best status positions. Based on calculations using more or less arbitrary numbers to operationalize these assumptions, Boudon concluded that if the educational system expands (that is, if more students survive through the higher regions of the educational system), the inequality of educational opportunity will steadily decline. This decline, however, is not followed by an according decline in the inequality of status opportunity. While the lower classes in Boudon’s artificial world are approaching the higher class in terms of educational attainment, there is no change in mobility, even if education is the only, if imperfect, determinant of occupational status.

A theory that is closely related to the one described above is the theory labeled the “labour queue theory”, by economist Lester Thurow (1975). Like Boudon, Thurow assumes that employers sort candidates for job positions according to their level of formal education, and other characteristics being equal, choose those who have the highest level of education for the best jobs. The reason for this is that employers seek to minimize the costs of on-the-job training, which is considered to be necessary for every job, and especially for those jobs that are high in the status hierarchy. A high level of education is then used as an indication that a candidate can be trained easily and efficiently. This means that those who are at the head of the labour queue will get selected for the best status positions, and that those who are next in terms of educational attainment will then be first in line for the next best positions. According to this principle, that is indeed very similar to the mechanism of job distribution that is applied by Boudon, a more equal distribution of educational attainment will not lead to more mobility, but instead to a phenomenon quite appropriately called “credential inflation”: if more people reach higher levels of education, the value of a degree will drop because it will still be only those with the highest levels that will be selected for the best positions. Empirical applications of this theory have indeed found this effect (see for example Wolbers, 1998). Although the labour queue theory seems to be cited a little more often then Boudon’s theory, and is in some respects somewhat more sophisticated, it has however the disadvantage that it does only model the transition from the educational system to the labour market, and does not consider the consequences for the relation between social background and occupational status. Therefore the labor queue theory will not be discussed in much detail in this study.

As I said before, it is not fair to say that there has been no theoretical progress in “mainstream” stratification research whatsoever since the days of Blau and Duncan. The work of Mare is of particular importance here. In a milestone article Mare (1981) reconceptualized the process of the generation of educational inequality as resulting from two independent processes. On the one hand, there is the mechanism of allocation of educational assets to students, that leads to a certain association between social background and success in the educational career. On the other hand, there
is the process of educational expansion consisting of higher numbers of pupils reaching higher levels of education over time. Mare maintains that the linear regression coefficient of educational attainment on social background - the traditional measure of educational inequality - will decline in times of educational expansion while the actual mechanisms of allocation remain unchanged. Instead, he proposes a different framework of analysis that (like in Boudon’s approach) conceptualizes the educational system as a series of transitions. By using logistic regression techniques, it is then possible to establish the effect of social background on the probability of surviving a transition, which provides a measure of educational inequality that is not influenced by educational expansion. Moreover, Mare shows that educational expansion does also independently influence the relation between social background and learning skills, by influencing the selection on these characteristics that takes place at each transition.

Because of the importance of these findings for contemporary stratification research, and the resemblance to Boudon’s approach to educational inequality, Mare’s theory will by considered alongside Boudon’s theory throughout this study. These theories and related arguments will be discussed in the next chapter, along with the empirical research connected to it.

1.3 Stratification and simulation

In his study Boudon used a method that is somewhat unusual in sociology: that of simulation. Loosely based on empirical evidence, a numerical example of what a stratification process in a modern society could look like was made up, consisting of “reasonable” axioms. Boudon’s conclusions were ultimately derived from his observations of this imaginary society. The feature of his study is probably one of the main reasons why it received so much criticism: many found that conclusions based on such a simple numerical example could not be justified. Although many of the counterarguments appeared to be not that strong (as shall be discussed in the next chapter), it is true that far-reaching conclusions based on just one numerical example are not very convincing.

Since the seventies however computational resources have developed very far, and today it is possible to apply simulation techniques that are much more sophisticated than the basically pen-and-paper method used by Boudon. For instance, it is possible to simulate not just one, but thousands of imaginary societies, and to study their stratification outcomes as related to their inputs. As will be argued at the end of the next chapter, this type of computer simulation allows for a better understanding of Boudon’s theory, and because of that also for a better opportunity to test the theory. Next to the direct interest in stratification processes, to examine the usefulness of computer simulation in stratification research is the second aim of this thesis.
1.4 Overview

In the next chapter (two) Boudon’s theory is discussed in more detail, and connections with other related theories are explored. I also give an overview of previous empirical research related to Boudon’s theory. Based on the conclusion of that chapter, I formulate research questions aimed at the further analyses of the models in chapter three, and develop a simulation model with the purpose of finding answers to these research questions. Chapter four presents the results of the simulation, and formulates hypotheses based on this results as answers to the research questions of chapter three. In chapter five an attempt is made to test these hypotheses empirically using data on school- and labor market cohorts, and the last chapter summarises the efforts and discusses the consequences of the results in a broader perspective.
2. Theories & Research

This chapter seeks to discuss theoretical efforts to model stratification processes, with a strong focus on the theory formulated by Boudon. However, related theories and arguments will also be discussed. Secondly, I will try to provide an overview of empirical work related to these models. At the end of the chapter I will draw up the balance about the usefulness of Boudon’s theory, and to point out the direction the research in this thesis is to take.

2.1 Boudon’s theory of social stratification

The aim of the following section in first place is to provide an overview of Boudon’s model of social stratification as presented in his book “Education, opportunity ad social inequality” (1974), and it’s implications.

In the presentation of the theory I will follow Boudon’s axiomatic approach. The theory is formulated of basic (general) axioms, that are complemented by auxiliary axioms that allow for calculations with a numeric example. These auxiliary axioms are typically the axioms that are varied in a simulation approach. The numbering of the axioms is copied from Boudon’s.

The theory actually consists of two separate models, that together describe the mechanisms that lead from social origin to social destination (status), via education. The first part is concerned with *inequality of educational opportunity* (IEO hereafter) or the transition from social origin to a certain level of education. The second part is concerned with the *inequality of status opportunity* (ISO), or status attainment. The IEO model consists of the following axioms:

B.1 School achievement (in the sense of learning abilities, or skills) is a function of social origin. The higher the social status of parents, the better the school achievement of a pupil will be on average (which does not mean that there are no high status pupils with low school achievement). This is what Boudon calls the primary effect of stratification.

B.2 School achievement (along with other variables that characterize an individual) does not change over time.

B.3 The educational system consists of a number of transitions, or branch points. At each branch point, a student will have to choose between the “higher curriculum” and continue his educational career, and leaving school. Returning to school after leaving is not possible, although this is a possibility in most existing educational systems.
B.4 At every transition, there is a probability $p$ that a student will choose the higher curriculum and stay in school. This probability is a function of both school achievement and social background, and doesn’t change during the educational career.

The direct effect of social origin on the survival probability $p$ is what Boudon calls the secondary effect of stratification, meaning that even when two pupils with different backgrounds have similar skills, the pupil with the higher background will have a higher probability of continuing in school. The idea behind this was first formulated by Keller and Zavalloni (1964). Their central argument is that reaching a higher educational level has a higher utility to students with a high background than to students with a low background, because high status students may lose status if they stop studying at a certain point while low status students might at this point keep the status they already had. Moreover, the relative costs of continuing might be higher for low status pupils, not only in a material sense but also in terms of social costs; they might for example lose friends when they choose to stay in school while their friends do not.

This decision model – which is implicitly a rational choice model - has been further developed by Breen and Goldthorpe (1997), who provided a formal rational choice model for the choice between further schooling or leaving the system. According to these authors, there are three main mechanisms that create class differences in the survival probabilities. The most important mechanism is relative risk-aversion: actors (parents or students) find not losing status more important than winning status, because they want to avoid ending up in a position that is worse than the starting position. For middle class pupils this is relatively easier than for high class pupils (because it is easier to reach the middle class than the upper class); therefore the latter are more likely to invest more in education. This corresponds to Boudon’s secondary effect, just like the second mechanism, being differences in resources. (The third mechanism is differences in achievement level, the primary effect). Because the secondary effect is the result of a cost-benefit analysis by the actors, it can be interpreted as the “rational” aspect of differences in survival probabilities, while the primary effect (which reflects differences in learning skills between social classes) might be interpreted as a more cultural aspect.

The study by Breen and Goldthorpe shows that it is possible to specify a more precise decision model as a basis for a formal theory of stratification. I will however not use this decision model here, mainly because it would be very hard to estimate its parameters empirically.

B.5 There is educational expansion; because of undefined exogenous causes, the survival probabilities at each transition increase over cohorts. Because some student’s survival probabilities are already very high, it is assumed that there is a ceiling effect at work, according to which high survival probabilities increase at a slower rate than lower probabilities. This ceiling effect is (as an auxiliary axiom) defined in the following formula:
\[ p_{t+1} = p_t + (1 - p_t)a \quad \text{where } (0 < a < 1) \quad \text{(equation 2.1)} \]

In this formula, some quantity is added to the survival probability \( p_t \) that is smaller the higher \( p_t \) (which constitutes the ceiling effect) and that is higher, the higher \( a \) (which governs the amount of educational expansion).

These five assumptions basically make up the IEO-model. To allow for computations, Boudon then introduces several more auxiliary axioms to give the model a quantitative form: he chooses three status levels, eight branch points, and concrete distributions of achievement levels and survival probabilities such that there is a direct and an indirect effect of social origin on the probabilities. He then uses this numerical example, designed as an ideal typical industrial society, to derive conclusions about IEO and its development. His main conclusions on the creation of IEO are that over transitions, the primary effects “die out” while the importance of the secondary effect increases; in other words the “cultural” effects of stratification are much less important than the “rational” aspect. On the development of IEO the most important conclusions are that IEO (measured as the ratios of the chances of children from different classes to reach high educational levels) declines steadily over time, while attendance rates increase at all levels.

The second part of the model concerns inequality of occupational opportunity. Again, I shall discuss the basic axioms of the model.

C1. All axioms of the IEO model hold.

This means that the process described in the ISO-model can be seen as the continuation of the stratification process started by the IEO-model; the simulated individuals that reached a certain educational level in the IEO-model now continue their career on the labour market. Thus the output of the IEO-model serves as the input of the ISO-model.

C2. It is assumed that the individuals of one cohort that entered the educational system together, will also enter the labor market at the same point in time.

This assumption is obviously unrealistic, as in reality the members of one cohort who leave the educational system early will because of that also enter the labor market before their fellow cohort members who chose a higher education will; achieving a higher level of education just takes longer. As we shall see later, this assumption creates some complications in the empirical testing of the
theory. However, modeling the process such that cohorts overlap as they do in reality would be very complicated. Therefore I will maintain this assumption.

C3. The simulated society is a meritocratic society, meaning that the allocation of status positions is mainly based on educational achievement. The members of this society compete for the best status positions, and the candidates with the best educational credentials have the higher chances. This assumption is made concrete by the introduction of the meritocratic parameter, representing the importance of education in the distribution process. The distribution process is as follows: Starting with the highest status positions and the highest educational level, the number of positions is compared with the number of candidates with the highest education. If there are more open positions than candidates, a certain percentage of the candidates will get a position. This percentage is the meritocratic parameter (0.7 in Boudon’s numerical example). If there are more candidates than positions, the meritocratic parameter is applied to the number of positions rather than to the number of candidates. This process is repeated until all positions are distributed among the candidates.1

This method of distribution has been much criticized (see for example Hauser 1979) for being unrealistic and hard to understand intuitively, and indeed in reality real world employers do not apply percentages to numbers of job applicants with certain educational credentials, or to the number of positions if there are too many candidates. Boudon’s own lack of justification for using this mechanism does also not contribute to the easy acceptance of the procedure. It is however possible to interpret the meritocratic parameter simply as the probability for a random actor with a certain amount of education to obtain the status position he is after (the highest still available). This can be easily seen in the case when the number of jobs is larger than the number of candidates: if for example there are 150 jobs for 100 candidates and the meritocratic parameter is 70%, then 70% of the candidates will get a position and the probability for any of the actors from this group to get a job is 0.7.

It is however less clear why the meritocratic parameter should be applied to the number of jobs in the cases when there are more candidates than positions: on first sight, this rule might seem rather strange. Also in this case however the parameter might be interpreted as the probability of obtaining the highest (still available) status position as a function. Consider the example that there are 200 candidates for 150 positions. There is now a shortage of jobs, which decreases the probability of getting a position regardless of the role of education: the probability of getting a job would be at most 150/200, even if a high education would be a full guarantee for getting a job. If a high level of education however only provides a 70% change for getting a position, the probability of obtaining a position becomes 150/200*0.7 = 0.525. This means that 105 candidates will get their desired positions,

1 This distribution process can be conducted in several ways; one could for example start with distributing all individuals with the highest educational level among the three status levels, then change to the next educational level, etc. Alternatively, one could first distribute all the highest status positions, then the second highest positions etc. This different methods do however not influence the results in any way.
which is indeed also the number which results when the meritocratic parameter of 70% is applied directly to the number of positions. Thus although Boudon’s procedure for the meritocratic distribution of positions is indeed unrealistic, it does provide a coherent model of a competitive process in which the meritocratic parameter reflects the probability of success as depending on education.

This “simple” version of the model is extended by adding so-called “dominance” effects to the model. This means that not only education is important in the distribution process, but also social origin, in order to provide a model that can account for the direct effect of origin on status that is usually found in empirical research. Education however still comes first: the highest status positions are distributed among those with the highest education, and among them, those with the highest social origin come first. Boudon implements this dominance effect by assigning different parameters of dominance to every combination of social class and level of education. Thus those with the highest level of education and the highest background have the largest dominance, and those with the lowest level of education and the lowest background have the smallest dominance.

The outcomes of this model are again analyzed, both for the simple version and the version with dominance effects, and for various points in time (following the IEO-model). Boudon’s most striking conclusion (from the simple model) is that while IEO declines steadily over cohorts (as measured by disparity ratios) and school attendance increases dramatically, social mobility (ISO) does not seem to be influenced by this process. The changes in mobility that do occur, do not follow any clear trend and appear to be random (which of course they are not, because the model is completely deterministic). On the other hand, even when IEO is high, there still is a considerable amount of mobility. Both these conclusions remain unchanged when dominance effects are introduced into the model.

The conclusion that mobility can be relatively stable while IEO changes provides, according to Boudon, an explanation to the well-known Lipset-Bendix paradox (which states that although industrial societies are very different with respect to IEO and organization of the educational system, they do not differ that much with respect to mobility, both geographically and over time; see Lipset & Bendix 1959). This can be attributed to the fact that while the structure of the educational distribution changes, the social class structure remains unchanged, which dampens the effect of a decrease in IEO. This argument is in line with the implications of Thurow’s job competition model, which also described a discrepancy in the development of both distributions. Related to this effect is the phenomenon of credential inflation: if the general level of education increases, the values of a certain educational degree decreases because it is only the relative educational position that matters in the labor queue (Thurow 1976, Wolbers 1998). This does not mean that actors do not have an incentive to
reach higher levels of education: on the contrary, actors will have more reason to strive for higher degrees to improve their competitive value. 2

Furthermore, Boudon explored the consequences of changes in the distributions of origin and status positions, though not in detail. His conjectures are that an increase in the proportion of high status positions would lead to more upward mobility, but would not change the earlier conclusion, while differences in fertility between the social classes would only have effect on IEO outcomes, but not on ISO outcomes.

The models as described above provoked much criticism and debate, like a devastating review in the American Journal of Sociology (Hauser 1976), a symposium in Social Science Information and several other studies devoted to Boudon’s book. In the following I will give an overview of the critical reactions on the work, and of reactions that aimed at further theoretical development of the model. Empirical research connected to Boudon’s theory will be discussed in the second part of this chapter.

The aspect of Boudon’s study that attracted most criticism was it’s methodology of simulation. According to many, the model was not realistic enough to allow for conclusions about real societies. Moreover, it was claimed that the conclusions would be only valid for this particular numerical example, and that some kind of “sensitivity analysis” would be necessary to allow for more general conclusions (Hauser 1976, Sørensen 1976).

Several aspects of the model were attacked as unrealistic. Among these are the assumption that survival probabilities remain constant over transitions and the assumption of the ceiling effect of educational expansion (according to Hauser 1976), there should not only be a ceiling effect, but also a “floor effect”). In relation to the ISO-model the meritocratic distribution principle, and the assumption that the social class structure does not change are considered unrealistic. Also, the analyses of the data that Boudon used to base his assumption on are thought to be biased, not precise enough or just wrong. Boudon’s reply to this criticism is that the model is not intended to be realistic; it is supposed to represent a typical modern society, and aims at explanation of processes rather then description (Boudon 1976). I think this is an important remark, and I agree with Boudon that it is not necessary for theoretical models to be “realistic”. Strictly speaking, theoretical models do not need to be based on observation at all, as long as they can be (and are not) falsified. However, as shall be shown later, there have not been many real attempts to falsify Boudon’s theory so far.

Besides these methodological issues, there is also disagreement about the way the results of the simulation should be interpreted. According to Hauser, the IEO model does not generate a decline in IEO when the difference in the mean level of education between social classes is considered, rather then Boudon’s “disparity ratio” of the probabilities of reaching a high educational level. And

\[2\] For a game-theoretic interpretation, see Elster 1976.
following this interpretation, it is of course no wonder that the ISO-model does not generate any differences in mobility, because there is no change in IEO.

In summary, the arguments against Boudon’s model, as most clearly formulated by Hauser, boil down to the following two points:

1. The assumptions of the model are not sufficiently supported by empirical evidence, and therefore unrealistic: this makes its conclusions invalid.
2. The conclusions do not always seem to follow from the assumptions, and it is unclear to what extent they depend on the choice of parameter values.

Of these arguments, the second is clearly purely theoretical. To some extend however that is also true for the first argument, for it implicitly claims that if the assumptions of the model were made more realistic, the conclusions could well be different. These theoretical issues were addressed in several studies that tried to overcome some of the problems of Boudon’s approach by using the formal methods of mathematical analysis. Fararo and Kosaka (1976) examined Boudon’s main conclusions regarding IEO using a formalized representation of the model, in which the “fundamental parameter” of the survival probability \( p \) representing the matrix of survival probabilities of the different subgroups takes a central place, but remains unspecified. Thus the mechanisms of the IEO-model are reformulated as functions of this generic parameter \( p \), and it is examined whether Boudon’s conclusions hold for all \( p \), only a limited range of \( p \), or perhaps only those values of \( p \) that Boudon used in his numerical example. The results of this analysis show that most of Boudon’s conclusion on IEO and it’s development appear to hold not only for the specific situation that he depicted by his choice of parameter values, but in most cases hold for \( p \) in general. It is shown for example that under the condition that \( p \) increases over time, attendance rates will always grow (one of Boudon’s conclusions) and IEO will eventually decrease. It is also shown that the educational process can be described in a continuous way (in contrast to the discrete representation of the original version with discrete social class, ability and time), without important consequences for the conclusions.

The analysis by Kosaka and Fararo thus shows that Boudon’s most important conclusions regarding IEO do indeed follow from the assumptions of the model. Many of these assumptions however were criticized for being unrealistic. Building on the analysis of Fararo and Kosaka, Raub (1984) addressed several of these more specific points, again applying mathematical methods that formalize the IEO model without specifying parameter values. This time however some of Boudon’s more restrictive – and possibly therefore unrealistic- assumptions are somewhat relaxed, in order to test the dependence of the conclusions on these assumptions. The first of these is the assumption that the survival probability does not change during the educational career; although survival probabilities differ between individuals according to background and ability, once established the same probability applies at every transition. Raub relaxes this assumption by allowing for different survival
probabilities per individual at each transition, as long as these probabilities are a function of the “fundamental parameter” $p$, such that the higher $p$, the higher the probabilities at each level. Another modification of the model concerns the way the survival probability changes over time. In the original model, $p$ increases over cohorts according to a ceiling effect defined as

$$p_{t+1} = p_t + (1 - p_t) a$$, where $0 < a < 1$  \hspace{1cm} \text{(equation 2.1)}$$

Raub adds a floor effect to this ceiling effect, using a logistic s-shaped function, which means that not only probabilities that were already high increase relatively slower, but also low probabilities increase relatively slower. This function is defined as

$$p_{t+1} = p_t + (1 - p_t) ap_t$$, where $0 < a < 1$  \hspace{1cm} \text{(equation 2.2)}$$

In fact this is precisely the “more realistic” way of educational expansion proposed by Hauser. The combination of these two modifications leads to four different versions of the model: the original version with fixed survival probabilities and the “generalized” version with differing survival probabilities, both with either a negative exponential or a logistic expansion function. All these four models can then be compared regarding the consequences for the creation and development of IEO, using Boudon’s conjectures as reference points. The results of this analysis show that also after making the model more general and perhaps more realistic, almost all of Boudon’s conclusions remain valid. The only conclusion that has to be modified is one about the speed of IEO decrease: the version of the model in which a logistic expansion function is added to the basic model shows a monotonous decline of IEO in time, while the basic model predicted a decline that would slow down in time. Both these studies have thus shown that Boudon’s model of IEO holds more generally than only for the numerical example that was applied in its original publication, and therefore that the second point of criticism as stated above – that the conclusions do not necessarily follow from the assumptions – should not be accepted too easily. Moreover, the study by Raub proves that it is not sufficient to simply claim that some assumptions of a model are unrealistic to dismiss the model. If making the model more complex in order to make it more realistic does not really change its outcomes, then why not stick to the simpler version? This means that the first argument against Boudon’s theory of IEO (and indeed against many “simple” models of society) that it is too simple or unrealistic should also be seriously reconsidered.
2.2 Expansion, transitions and selection: Mare

Another approach that also conceptualizes the educational system as a sequence of transitions, and has been something close to a breakthrough in stratification research, is that found in the work of Robert Mare. According to Mare, the process responsible for educational inequality actually consists of two different processes, that should be considered separately. In the first place, there is the mechanism that allocates educational assets to people with different background characteristics. This means that at every transition to a higher level of education, social background has an effect on the probability of surviving the transition, both through intervening (but usually unmeasured) variables like learning skills that are related to social background, and directly. In the second place, there is the independent process of educational expansion (part of more general modernization processes in society), which means that the number of students reaching higher levels of education increases over time. Mare shows that the standard way of representing educational inequality, namely the OLS linear regression coefficient of highest completed level of education on social background is not able to disentangle these two processes. When the educational system expands IEO measured in this way tends to decline, while the association between social background and the probability and surviving a transition remains stable. However, when the educational system expands but the association between social origin and surviving transitions becomes stronger, the two trends can offset each other such that IEO, expressed as a linear regression coefficient, remains stable over time. Regarding Boudon’s model, Mare concludes that although it’s conclusions regarding the relation between expansion and inequality may be correct, it fails to disentangle the two different processes. The decline of IEO results directly from the increase of the survival probabilities because of the ceiling effect Boudon assumes, but is not necessarily related by a changing effect of origin on school continuation (Mare 1981).

It should however be noted that although Mare’s analysis in itself might be valid, his judgement of Boudon’s theory might be somewhat misguided because he considers the model to be a statistical model. In the explanatory model, the two different processes (allocation and expansion) are conceptually perfectly disentangled: for the first cohort (in Boudon’s sequence of four) the primary and secondary effects are established, and in the other three cohorts there is only a certain amount of expansion, but no change in the size of the primary and secondary effects per se. Thus Boudon shows that IEO can decrease while primary and secondary effects remain stable, just because of educational expansion, which means that Boudon’s theory and Mare’s analyses do actually not contradict each other. In fact, one would expect the patterns described by Mare to be predicted by Boudon’s model. For this reason, the two models will be discussed together in this study.

Mare’s analyses have some other interesting implications regarding what Boudon calls the primary effect: the association between social origin and learning skills. If school continuation is at least partly dependent on skills, the sequence of transitions can be considered to be selecting on these skills: it is only possible to survive a certain transition if you have the right capabilities. As a result, heterogeneity on skills among students will decrease over transitions. This also means that educational inequality
(not controlled for differences in capabilities) will decrease over transitions, simply because students tend to differ less on variables that are related to social background higher in the educational system. However, if the educational system expands, the number of students at each transition will increase, thereby increasing the heterogeneity on skills among the students at a certain transition. The result is that the association between social background and school continuation at a certain transition will become stronger if the educational system expands. Again, it can be expected that if Boudon’s IEO model does indeed provide a real explanation of stratification patterns, then these effects should also be produced by Boudon’s model.

Although most attention has been drawn towards the IEO-model, more or less similar efforts of examining the robustness of Boudon’s claims have been directed at the ISO part of the theory more recently (Müller-Benedict 1999). Instead of using mathematical methods of deriving conclusions for unspecified parameter values, Müller-Benedict applied computer simulation techniques to examine the behavior of the ISO-model. This method, that is similar to the method of simulation that will be used later in this study, consists of varying the parameters of the model over a relatively large range of possible values simultaneously, and by evaluating the sensitivity of the outcomes to this variation. For example, where Boudon draws his conclusions on stability of mobility from a situation with a level of meritocracy of 70%, this type of simulation analysis would typically test whether the conclusions also hold for a level of 80%, 90% and so on, while at the same time also varying the other parameters of the model. Although this method is less exact than mathematical analysis, it is capable of providing an understanding of the robustness of a model where mathematical representation is very complicated (or sometimes even impossible) or hard to grasp for those who are unfamiliar with mathematical analysis. In his simulation Müller-Benedict varies three parameters of the ISO model: the level of IEO (defined as an odds’ ratio; Boudon used “disparity ratios”), the level of meritocracy and the level of dominance. The distribution of available status positions is still assumed to be equal to the distribution of social background. Both the simple version of the model (without dominance) and the extended version (with dominance) are considered. The analysis of the simple model shows that Boudon’s important conclusion that the level of IEO does not substantially influence mobility does not hold generally: although ISO (also expressed as an odds’ ratio) indeed remains more or less constant in the range of IEO that Boudon considered, it does decline when lower levels of educational inequality are reached. Furthermore, the analysis shows that ISO increases when the level of meritocracy increases. The analysis of a more complex version, in which also a parameter of dominance operates that is supposed to be equal to the meritocratic parameter, reveals a much more complex pattern of ISO development. More meritocracy still leads to less mobility, but the relation between educational inequality and mobility now shows a rough U-shaped pattern, with many exceptions in between. This complex structure can be explained by the fact that as long as enough high status positions are available to cater a group of people with a certain educational degree, the extended model behaves like the simple
model, but when the limits of the availability of positions are approached, social background also starts to play a role. This means that again some kind of credential inflation takes place: the more people reach a certain level of education, the smaller the relative importance of having this degree for obtaining a high status position becomes in comparison with other characteristics.

### 2.3 Empirical efforts

In his book, Boudon himself already indicates that empirical testing of his theory would not be easy. One problem is that many of the parameter values used by the model seem hard to measure; the “meritocratic parameter” would be a good example, but also less exotic parameters like the association between social class and learning capabilities (the primary effect) are not frequently found in empirical data. Another problem might be that the scope of the theory is quite large: it contains both statements about the way individual families take decisions in education and about macro level mobility structures. Maybe for these reasons, the number of empirical applications of Boudon’s theory is relatively small (compared with for example the number of replications of Blau & Duncan’s study), and of those most studies are limited to the testing of Boudon’s more general conclusions. In this section I will provide an overview of studies that have tested different aspects of Boudon’s theory.

With regard to IEO, most studies have concentrated on the mechanisms of primary and secondary effects. Bosker, van der Velde and Otten (1989) for instance compare predictions from different theories on the creation of IEO, including theories by Kohn, Bernstein, Bourdieu (all of which state that differences in educational attainment are produced by “cultural” differences) and Boudon. Most attention is paid to Kohn’s theory, which states that differences in “self-determination”, as transferred by parents, cause differences in educational attainment between social classes. The different theories are compared using Dutch longitudinal data collected in the 1970s. Educational level is measured both at the end of primary school and at the end of secondary school. Intelligence, parents’ value orientations and parents’ occupational status are also measured. Regression analysis shows that in accordance with Boudon’s theory, there are indeed primary and secondary effects: intelligence has an effect on educational level at the end of secondary education, but there is also a substantial direct effect of social origin. Kohn’s theory however is not supported: although the higher strata does indeed display a higher level of “self-determination”, there is no significant effect of the level of self-determination on educational attainment.

Boudon claims that the strength of primary effects declines during the educational career, while the secondary effects become stronger, and are in the end stronger than the primary effects. Based on a comparison of correlation coefficients however, Bosker et al find that, at the end of the educational career, primary effects are still stronger than secondary effects. The authors thus conclude that
although Boudon’s theory seems to perform a little better then the competing theory by Kohn, it can also not be convincingly confirmed.

The conclusion that Boudon’s model tends to underestimate primary effects is also found in research on more or less the same Dutch school cohorts (Dronkers et al, 1982). In this research, the model is tested by attempting to adjust it to the Dutch situation. Although the existence of primary and secondary effects is again confirmed, the pattern of survival probabilities found in this research initially show that learning skills are in general more influential than social background, which only plays a role for those students with high achievement levels. However, when the analysis is adjusted for changes in the Dutch educational system in the 1970s (which results in a different definition of transitions), the predictions of the theory appear to fit the data somewhat better: the secondary effect now plays a more important role (p. 98, 112). The authors conclude that the fit of the model to earlier cohorts is not that good because of the very strong differentiation of the Dutch educational system at the time: the most important decision by far had to be taken at the first branch point. In later cohorts however, the sequence of transitions became more equal, resulting in better fit of the model.

The hypothesis on primary and secondary effects is tested in a different way by Halsey, Heath and Ridge in their study *Origins and Destinations* of the British school system (Halsey et al, 1980). These authors state that, if Boudon’s hypothesis of exponential secondary effects is correct, differences in survival probabilities between social classes should become larger later in the educational career. The results of analyses of the educational careers of English students do however not confirm this expectation: differences in survival probabilities between social classes seem to become smaller rather then larger. According to the authors this suggests that the importance of secondary effects decreases instead of increases: at the end of their educational career, the chances of low class students are almost equal to those of higher class students. The hypothesis is tested in still another way by creating fictitious data, based on the original data. In one dataset, it is assumed that only primary effects play a role; in the other, that only secondary effects exist. It is then concluded that the dataset with only primary effects resemble the real data best, and therefore that primary effects are more important then secondary effects. The explanation for the discrepancy between these findings and Boudon’s hypothesis is, according to the authors, that the number of transitions (branch points) in the British educational system is much lower than that in Boudon’s simulation. The consequence is that secondary effects do not become as strong as they do in Boudon’s findings. Another reason is that the British educational system aims at an early selection, which reduces competition between pupils after selection. According to Halsey et al, this also reduces the size of secondary effects.

The ISO model has been subject to a number of empirical studies as well. Some of these are concerned with the more general idea of education as a relative good, that is not only an ingredient of Boudon’s theory but also of for instance Thurow’s “labor queue” model. In order to test the latter theory,
Wolbers (1998) analyses Dutch labor market cohorts and finds that the returns to education have
decreased between 1911 and 1991; in other words there has been credential inflation. According to
Wolbers, the reason for this is that the educational system has been expanding at a faster rate then the
structure of status positions. This is a confirmation of Thurow’s theory, but also of other job
competition theories, that consider education to be a relative good.

An other study that aims at testing this type of theories was conducted by Ultee (1978). He compares
“utilitarian” theories that consider education to be a “positional good” (Boudon, Thurow) with theories
that consider education to have an absolute value. According to Ultee, the theories by Boudon and
Thurow imply that because of credential inflation the association between the relative educational
level (relative to the distribution in society) and occupational status will remain stable. The intercept of
a linear regression model of occupational status on absolute educational level is supposed to decrease
over time, because people with a certain level of education will be forced into lower status positions as
the number of people with higher levels of education increases. Moreover, the utilitarian theories
would predict a higher squared correlation coefficient between relative level of education and status
then between absolute education and status. “Revisionist” theories on the other hand, assuming an
absolute value of education, would predict that regression coefficients increase over time, while the
intercept declines.

Analysis of American mobility data over a period of twelve years (1963-1975) shows that the
utilitarian predictions are closer to the observed trends in the relation between education and status
then the revisionist theories: regression coefficients tend to be stable, while the intercept for absolute
education declines. However the utilitarian predication that the (squared) correlation coefficient would
be higher for relative education is not confirmed. None the less, Ultee concludes that the utilitarian
theories perform surprisingly well.

Job competition models predict credential inflation as a result of educational expansion, because the
development of the structure of status positions does not keep pace with the changing distribution of
education. If this is true, then credential inflation should not occur when the distribution of education
is artificially kept in line with labor market demand. Andorka (1976) shows that this was indeed the
case in socialist Hungary, where both the educational system and the job structure were under state
control, and where one could be adapted to the other. This results in no credential inflation, and an
increase in mobility for women. Andorka takes this to be a confirmation of Boudon’s theory, while it
is in fact just not a falsification, showing that the theory goes wrong where it should go wrong.

While the empirical studies discussed so far in relation to the ISO-model mostly look at credential
inflation, there are some studies that have paid more attention to ISO itself, that is the relation between
social origin and status. Again, approaches and conclusions are very different. Müller-Benedict, in his
simulation study already discussed before, found when education expands, mobility will initially

3 Ultee is referring to revisionist Marxist theories, which claim that education is a means of power, and therefore
has an absolute value.
increase but at a certain point reach a saturation level (caused by the fact that the status distribution
does not change) and than increase again when IEO further declines (this conclusion are derived from
the simulation model that uses dominance effects in addition to “meritocratic” effects). This theoretical
finding was roughly (qualitatively) tested by looking at German mobility trends. Given the fact that
IEO has declined in Germany during the twentieth century starting at a relatively high level, it is
predicted that ISO (measured in odds ratios) will show a steady but slowing decline, approaching the
maximum level of mobility. This indeed appears to be the case, and this is taken as a confirmation of
Boudon’s models. Other studies however reject Boudon’s ISO theory based on the finding that a high
level of education really can be a source of mobility (Delmotte and Despriet 1984).

2.4 Conclusion
The first part of this chapter discussed Boudon’s theoretical models of IEO and ISO,
counterarguments against these models and further analyses, as well as some related theories. From
this overview I think it is justified to conclude that Boudon has indeed formulated a theory that really
models (rather than describes) processes of social stratification by specifying the mechanisms that
operate in these processes. Moreover, although Boudon’s methods of using a numerical example were
not that convincing, additional analyses have especially in the case of the IEO model shown that
Boudon’s conclusions are not very much dependent on arbitrary parameter values or simplifying
assumptions; in other words the IEO model seems to be quite robust. This is less the case for the ISO
model, which has received much less attention. Here it was shown that Boudon’s most important
conclusion of stable mobility does not hold within a broader parameter range.
It also appeared that Boudon’s IEO theory relates to Mare’s influential analysis of stratification
patterns and measures in interesting ways. The exact nature of these relations is however still
something to be examined.
It is much harder to distill conclusions from the discussion of empirical efforts related to Boudon’s
theories. Both the methods applied and the conclusions derived are very different. With regard to IEO,
the general conclusion seems to be that although primary and secondary effects exist, they do not
always have the effects predicted by the theory. With regard to ISO, credential inflation is often found,
but other consequences of the model are hardly investigated.
However different, most empirical studies discussed to me seem to have one common characteristic:
they only test Boudon’s general conclusions, like the hypothesis that secondary effects become
stronger than primary effects, or that educational expansion does not lead to more mobility. They do
not really apply the theory to the situation is supposed to be tested on. A good example is Halsey’s
conclusion that primary effects remain stronger then they do in Boudon’s study, supposedly because
the number of transitions is lower in the British educational system. In my opinion, this is not a real
application of the theory. A real test of the IEO model would in this case be to see what the model
predicts in a situation with a low number of branch points, and compare that with the existing situation. From this point of view, Boudon’s numerical example is only a special case of a more general theory, that specifies how certain inputs lead to certain outputs. In most empirical contributions however only the outputs of the model that results from one set of inputs (the numerical example) are considered. Maybe this is partly a result of Boudon’s own assertions that the models should not be considered as representing any existing society, but the ideal type society. However I think that the theory could only be put to the test by examining how it predicts different outcomes in different situations. Thus my general conclusion of this chapter would be that although theoretically Boudon’s theory seems to be promising, it has never been tested sufficiently. I propose to test it by trying to figure out how the outcomes of the model change if it is applied to different inputs, and by comparing the predicted outcomes to existing situations that differ accordingly. The first next step will be to derive more accurate hypotheses from Boudon’s stratification theory by means of computer simulation, as will be explained in the next chapter(s).
3. Research Questions & Simulation Design

In the previous chapter I concluded that most of the previous research concerned with Boudon’s stratification models failed to relate the theory properly to the empirical findings that the theory was supposed to be tested to. More specifically, in most cases only the general conclusions that Boudon derived from his single numerical example were compared to the outcomes of a specific situation. In my opinion, this method does not tell us much about the real value of the theory, because the numerical example that Boudon used in his theory should only be regarded as a special case of a more general theory that can be described by the few relatively simple assumptions that I discussed at the beginning of the previous chapter. The theory can be tested by applying it to a specific existing situation, but applying means that the theory should be adapted to this situation to see what it would predict there. For instance, a test of the predictions concerning the developments of primary and secondary effects over the educational career is not very informative if the predictions from Boudon’s example with eight branch points are compared with the outcomes in the British educational system, that has much fewer branch points (see Halsey et al., 1980). A better test in this case would be to work out the theory’s predictions for the appropriate number of branch points and compare those predictions to the data.

Instead, I propose to further explore the model’s consequences, and see how the outcomes change when the parameters of the model are varied. One way of doing this is using computer simulation to systematically vary parameter values and compute outcomes, and that is the approach that will be used here. Another way would be further mathematical analysis, comparable to the method applied by Fararo and Kosaka, and Raub. Although that method would be more precise, computer simulation has the advantage of being relatively easy to apply, while it can also be used in cases that are so complex that mathematical analysis would not be very successful (see for example Buskens 1999).

The results of the simulation will consist of a collection of variants of Boudon’s original stratification models. Statistical methods can then be used on this dataset to generate hypotheses to be tested with real data.

In the following sections I will formulate a number of research questions that will guide the simulation design and analysis. It is important to keep in mind that simulation is a means of theory construction and not empirical research, although it may sometimes look like it.

3.1 Research questions

As explained above, the aim of the analysis described in the next sections is to investigate how the outcomes of the model depend on the values of the parameters. For this purpose, several research questions will be formulated concerning the effects of the model parameters on outcomes like IEO, IEO development and ISO. In some cases, it is also possible to formulate certain expectations on these
relations, based on former research and “common sense”. One should realize that these expectations are not really hypotheses in the normal sense of the word, since they are not expectations about real phenomena but about the theoretical implications of a theoretical model. It is these implications that are the real hypotheses.

_Inequality of Educational Opportunity_

The first part of the model describes how individuals with certain characteristics arrive at a certain educational level. The collective result of this process then serves as the input for the second part of the model, where individuals compete for status positions. It seems sensible to analyze these two models separately, and to consider educational inequality as a dependent variable first. There are four parameters in the model that might influence this variable. First, there are the two parameters that together determine the influence of parental status on the survival probabilities. According to the theory, stratification has both a primary and a secondary effect: parental status has an effect on the pupil’s school achievement level which in turn has it’s effect on the survival probabilities (the primary effect), while parental status also has an independent effect on the survival probabilities (the secondary effect). In the simulation program, both the influence of parental status on achievement level and the influence of parental status on the survival probabilities are varied, and can be expressed as regression coefficients. It can be expected that IEO will increase when the influence of these two parameters on the survival parameters increases, simply because in that case the influence of social origin on school continuation increases. However it is not a priori clear if, for instance, this relation is linear. Moreover, to get a better understanding of the workings of Boudon’s theory it is helpful to get some “feeling” for the quantitative sizes of these kinds of effects. This results in two research questions:

_Retal question 1:_

*What is the influence of the association between parental status and achievement level on IEO?*

_Retal question 2:_

*What is the influence of the association between parental status and the survival probability on IEO, taking the primary effect into account?*_

Besides these mechanisms on the individual level there is also the structure of the educational system, implemented in the model as the number of branch points. In the original version this number is more

---

4 In theory, one can make a difference between school achievement and skills: the latter can be defined as the real abilities of a student, while the former can be defined as a measure (for example by test scores) of these abilities. Since I am not interested here in measurement techniques however, I will ignore this difference: both “skills” and “school achievement” refer to the real learning abilities of a student.
or less arbitrary set to eight, but as Halsey et al indicated, the number of branching points (or educational levels) can be expected to have it’s effect on IEO (Halsey et al. 1980). Therefore:

**Research question 3:**

What is the influence of the number of educational levels on IEO?

The three variables described above influence IEO directly for a given cohort. But as described before, the model also includes a mechanism of educational expansion, which lets all survival probabilities increase over cohorts, while applying a ceiling effect. According to Boudon, this would lead to a decline in IEO, and this conclusion was confirmed by later mathematical analyses. Mare (1981) also concluded that educational expansion would result in a decline in IEO, when expressed as the linear relation between origin and highest level completed. The fourth research question aims at exploring the relation between educational expansion and IEO more precisely:

**Research question 4:**

What is the influence of educational expansion on IEO?

So far only the independent effects of the various variables were considered. It may however be expected that the sizes of the effects of the parameters depend on the value of the other parameters. For instance, one could expect that the primary and secondary effects have more impact the more educational levels the system contains. The next research question considers these kind of interaction effects:

**Research question 5:** To what extend do the effects of the number of educational levels, the primary and secondary effect, and educational expansion interact?

*Selection and IEO*

The research questions described above all consider IEO as an end result, when all the members of the cohort have completed their educational career. It is however possible to have a closer look at what happens at the different transitions. As Mare indicated, the overall effect of social origin (parental status) is the results of two different mechanisms. (Mare 1981). On one hand, the mechanism of allocation of education to persons (via skills) determines the survival probabilities. A decrease of this effect would lead to a weaker effect of social origin on education. On the other hand, educational expansion leads to a larger variation in individual skills, leading in turn to a higher overall effect of parental status. This means that the statistical effect of social origin on education can remain stable
while the underlying mechanism are changing, in opposite directions. With the help of simulation, it is possible to disentangle these different effects. 

Another implication of Mare’s analysis is that one should see the statistical effect of social origin on achievement at a certain educational level increase as a result of educational expansion, because the variation on achievement increases. I investigate this question for the highest level:

*Research question 6: What is the effect of educational expansion on the association between social origin and achievement at the highest educational level?*

**Inequality of Status Opportunity**

While the variables described above lead to a certain distribution of education and a certain level of IEO, they do not (at least not in this simple version of the model, where dominance effects are not taken into account) directly influence the second step in the model, the transition from a certain educational level to a status position. However there are other parameters that have their influence in this part of the process.

The first variable of interest is endogenous, namely IEO. One of Boudon’s most interesting claims is that mobility can remain stable despite a decrease in IEO, but he does not offer any form of sensibility analysis of this claim. The fifth research question addresses this issue. Müller-Benedict (1999) looked into this problem using a simulation approach, and concluded that Boudon’s conclusion does not hold below a certain level of IEO. The research question here can be considered as an attempt to replicate this result.

*Research question 7: how does the level of ISO depend on the level of IEO?*

The same can be said about the influence of the meritocratic parameter, which stands for the degree to which education determines status. Müller Benedict found that ISO increases with the meritocratic parameter (that is, in the simple version of the model without dominance effects). Again, I’ll try to replicate this result.

*Research question 8: how does the level of ISO depend on the value of the meritocratic parameter?*

In this study I will only consider the “simple” version of the model, where only education has an effect on status and further dominance effects of social class play no role, although dominance effects (i.e. the direct effect of social origin on status, controlled for effects of education) are usually found in empirical research (see for example Blau and Duncan, 1967). One reason for this is that the inclusion
of dominance effects makes the model more descriptive and less explanatory: it just adds another parameter that defines a relation, but does not explain it. This is especially true for the way Boudon implemented dominance effects: by more or less arbitrary assigning dominance coefficients to social groups (Boudon 1974, p. 155). In the interpretation of Müller-Benedict, there is a dominance parameter that is equal to the meritocratic parameter, but there are no substantial arguments why this should be the case. In my opinion, it is at this point more interesting to see how far we can get by applying the simpler and “cleaner” meritocracy model. Another reason is that the parameter for dominance would be very hard to establish empirically. This is already true for the meritocratic parameter, but here it is still possible to find reasons why the relation between education and status could be sometimes stronger or weaker (for example, there seems to be a connection with labor demand (Wolbers 1998)). For dominance, this is even harder.

Is his analysis of Boudon’s model, Müller-Benedict maintains the assumption that the distribution of status positions does not change between generations, i.e. the structure of available positions is the same as the distribution of parental status. Most authors however would agree that the status distribution does not remain stable, but changes as a result of “modernization processes” (Wolbers 1998): the idea is that the share of higher status positions would increase because of the rising average level of education, technological innovation, etc. Therefore it makes sense to investigate the influence of a change in the distribution of status positions:

Research question 9: What is the influence of change in the distribution of status positions on ISO?

Note that Boudon’s original theory can not provide any hypotheses on this subject, because it assumes the job structure to be invariable. This means that in the simulation the model will have to be adapted to investigate this question.

As with the research questions on IEO, one can again ask whether the effects of these three variables work independently, or interact with each other.

Research question 10: To what extend do the effects of IEO, the meritocratic parameter and the status distribution interact?

The effect of education

A final question of interest concerns the effect of education on status. How does the structure of the process of job competition affect the “returns of education”? The dependent variable here would be the regression coefficient of status on education; the independents variables would again be the three variables discussed before. Thus the research question reads:
Research question 11: How is the effect of education on occupational status influenced by the level of meritocracy, IEO and the distribution of status positions?

Something should be said here about the role of the meritocratic parameter in this question. The meritocratic parameter can in general be described as the importance of education in the selection process in the labor market. To investigate the influence of the value of this parameter on the effect of education then seems tautological, or at least not very interesting. It should be remembered however that the value of the meritocratic parameter does not directly determine the effect of education, since the process of status attainment is a competitive process: the importance of an educational degree in this process (as for instance in the value that employers attach to education when they choose between job applicants) should be distinguished from the statistical effect of education on status, which is the outcome of the process.

3.2 Simulation design

The research questions outlined above will first be examined with the help of a computer simulation model, with the aim of deriving hypotheses that can then be tested on real data. In remainder of this chapter I explain the design of the simulation program. In the next chapter I discuss the simulation results, and formulate hypotheses.

In order to answer the questions stated above, I will make use of a computer program that calculates stratification outcomes according to Boudon’s theory, using different inputs. More concrete, the program reproduces Boudon’s numerical example repeatedly while systematically varying the model parameters that are relevant for the research questions: the strength of the primary and secondary effects, the number of transitions or branch points, the extend of educational expansion, the degree of meritocracy and the change of the occupational structure. Each combination of parameter values constitutes a different virtual stratification system, and the final results of the simulation is a dataset with information on inputs and outputs of a large number of simulated stratification processes, that can then be analyzed.

The simulation model used here follows Boudon’s original model closely. There is however one big difference concerning the procedure: instead of calculating the proportions of groups, like Boudon does, this model works with individual cases. In Boudon’s version there are only proportions; in my version there really is an individual case for every virtual member of the groups. This difference has some consequences, which will be explained later.
The model starts with a group of 10,000 simulated individuals, of whom 1000 are from the highest status group (their status is “3”), 6000 are from the lowest status groups (status = 1) and 3000 are in the middle (status = 2). These people are then all awarded some level of school achievement that is positively correlated with status; this is what Boudon calls the “primary” effect of stratification. The positive correlation is created by a simple linear function in which a certain amount of “noise” is included. In Boudon’s version the correlation is about .2; in this version the correlation is determined by a parameter that controls the amount of noise:

\[ \text{Skills}_{\text{temp}} = \text{Status} + \text{noise} \div \beta \]  
(equation 3.1)

in which noise is a random variable uniformly distributed between -1 and 1, and \( \beta \) is a parameter varied in the simulation.

The values that result from the function in equation 3.1 are subsequently used to create a variable for achievement level the can assume the values 1, 2 and 3, with 3 being the highest achievement level. The next step is assigning the initial survival probabilities to the simulated individuals. According to the theory the survival probabilities are not only positively correlated with achievement level, but also with parental status directly (the so-called “secondary effect” of stratification). The survival probability is thus a function of both the level of school achievement and of status. Again this is basically a simple linear function, modified in such a way that the outcome is a value between 0 and 1, and also in such a way that the people with the lowest status level and lowest achievement level get a survival probability of .15, and those who score highest on both variables get a survival probability of .85. These values (almost) correspond to those used by Boudon. In this function, the size of the “secondary effect” is varied by changing the weight of parental status as compared to achievement level. The first step in the calculation of the survival probability is:

\[ p_{\text{temp}} = (\text{skills} + \gamma \times \text{status}) \left( \frac{(\gamma \times 3 + 3) - (\gamma + 1)}{2 \log(0.85 / 0.15)} - 2 \log(0.85 / 0.15) \right) \]  
(equation 3.2)

in which \( \gamma \) is a parameter varied in the simulation.

Subsequently, the survival probability is computed as:

\[ P = \frac{e^{p_{\text{temp}}}}{e^{p_{\text{temp}}} + 1} \]  
(equation 3.3)

Now that the 10,000 cases all have a survival probability, they can start their educational career. This is done as follows: starting at the lowest level of education, an individual can pass to the next level.
with a chance that equals his survival probability. If he passes to the next level, he again has a chance
to pass determined by his survival probability. If he does not pass, he has “left school” and is awarded
the educational level he passed last. This process is repeated for each individual until he drops out of
school or reaches the highest possible level.

This is where the procedure used here differs for the first time from the procedure used by Boudon. In
the original model, the proportion of a certain group that reaches educational level \( x \) is simply survival
probability to the power of \( x \); in this version the proportion is the result of a stochastic process. The
consequence is that where the original model is entirely deterministic, the version used here can result
in slightly different outcomes when repeated. The advantage of this method is that the output of the
simulation consists of datasets that look like real school cohorts, and can be analyzed accordingly.

After the simulated individuals have thus completed their educational career, they are distributed
among the available status positions, using Boudon’s ISO-model. It is assumed that the people of one
cohort enter the labor market simultaneously. The simple version of Boudon’s ISO-model, in which
only formal education plays a role, is used here. Starting with the highest educational level, the
number of candidates is compared with the number of available positions on the highest status level.
The structure of available positions is determined by another parameter, that governs the relative share
of high status positions in the distribution. In the original model, the distribution was the same as the
distribution of social origin; here the share of high positions can be changed by replacing a certain
number of low status positions by high status positions, resulting in an increase of the proportion of
high status positions. This reflects the idea that as societies modernize, the job distribution shifts
towards more high status positions.

If there are more positions then candidates, \( x\% \) of the candidates, where \( x \) is the level of the
meritocratic parameter, gets a position on this level. If there are more candidates then positions, the
parameter is applied to the number of positions, so \( x\% \) of the positions is allocated to the people with
the highest educational level. The rest of the people with the highest educational level then compete
for the second highest status level, etc. When all the cases in the cohort have received a status position,
several measures describing this cohort are calculated, among which regression coefficients to
measure the primary and secondary effects, IEO, ISO etc. The results are then saved and the
simulation round ends.

Like in the original version, the process is then repeated for a sequence of cohorts (Boudon uses four
cohorts), in which the survival probabilities of all groups increase gradually using the function

\[
p_{t+1} = p_t + (1 - p_t) a \quad \text{where} \quad 0 < a < 1
\]

(equation 2.1)

In Boudon’s version, \( a = 0.1 \); in the simulation this parameter is varied.
This whole process (including the stratification process for several cohorts) is then again repeated, while varying the parameters of the model. The parameters varying are the rate of educational expansion, the number of education levels, the influence of parental status on achievement (primary effect), the direct influence of status on the survival probabilities (secondary effect), the meritocratic parameter and the proportion of available high status positions. These parameters and the way they are varied in the simulation are described in table 3.1.

Table 3.1: description of parameters used in the simulation

<table>
<thead>
<tr>
<th>Simulation parameter</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Increment</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (expansion; see eq. 2.1)</td>
<td>0.1</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>β (primary effect; see eq. 3.1)</td>
<td>0.1</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>γ (secondary effect; see eq. 3.2, 3.3)</td>
<td>0.2</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>Number of educ. levels</td>
<td>2</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.5</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>Prop. of high status positions</td>
<td>0.1</td>
<td>0.28</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Apart from the model parameters described above, there are also other aspects of the models that one could consider of varying, but that are not varied here. Among these are the number of different status positions (now invariably three), the number of different achievement level (skills) values (now also three), or the way educational expansion is implemented (now with only a ceiling effect, but other ways are possible: see Raub 1984). Of course many more other possibilities are thinkable, but for this study have chosen to concentrate on those that relate to the eleven research questions described in this chapter.

In some cases, the model might not have a valid outcome. This happens when, for instance, after the people with a certain educational level have been distributed among the three status levels, there are still people with this educational level left that have not received a status. The same problem can occur when at a certain status level there are still places left but no people to fill them. Boudon apparently has not noticed this problem; nor has any other student of this model to my knowledge. In these cases, the program reports that there is a problem and returns “0” for the measures concerning the allocation of status positions (Which occurred in about 15% of the rounds).

The whole simulation process eventually results in a dataset of 84,000 different combinations of parameter values and corresponding outcomes. The analysis of these simulated data is the subject of the next chapter.
4. Simulation Results

The following section discusses the results of the analysis of the data generated by the simulation. These data consist of information on the 84,000 different “cohorts” that were produced by the program: parameter values and macro outcomes regarding IEO and ISO. The relations between inputs and outputs as described by the research questions in the previous chapter can now be investigated using regular statistical methods like regression analysis, where one simulated cohort constitutes one observation. In the following, first the choices of dependent and independent variables are discussed. The results are then presented, ordered by research question. The answers to the research questions are formulated as hypotheses.

4.1 Dependent variables

Inequality (IEO and ISO) can be expressed in many ways. One of the most common ways to express a level of inequality is the odds’ ratio. Müller-Benedict, for example, uses this measure in his simulation efforts on Boudon’s model. Boudon’s measure of IEO, the disparity ratio (the ratio of the probabilities of reaching a certain educational level between social classes), is closely related to the odds’ ratio. One of the reasons for using odds’ ratios is that they are easy to calculate and to interpret. A disadvantage however is that they appear to be very sensitive to outliers: a low frequency in one of the cells of a table that is used to calculate an odds’ ratio can result in an extremely high value. Moreover, if only one odds ratio is used to express the inequality in group, the measure is somewhat inaccurate because it only takes into account the frequencies in the corner of (for example) a mobility table, while ignoring the rest of the distribution. To create a more accurate picture of inequality a series of odds ratios would have to be used. To be able to use inequality as a dependent variable, we need a single measure.

The data produced by the simulation program allow for a more precise measure for inequality: because the simulation is based on “individual” data it is possible to calculate the regression coefficients of attained education and status on social origin for each combination of parameter values. These coefficients can then be used as dependent variables in the analysis of the simulation data (where one observation represents one such a combination), expressing the levels of IEO and ISO.

In the case of IEO, there is one complication. Because the number of educational levels is varied between cohorts, the scale of the variable “education” also changes. For instance, if the number of educational levels in a cohort is two, the possible values for attained education are 1 and 2, 2 being the highest possible education. However when the number of levels is three, the highest possible level of education rises from 2 to 3, influencing the regression coefficient of education on origin. To solve this problem, I use the standardized regression coefficient of education on origin. This coefficient does not depend on the scale of the variable “education”.


Another dependent variable that is considered in the analyses is the association between social origin and school achievement on the highest level, or the *conditional primary effect* (see research question 6). Here the regression coefficient of school achievement on social origin, estimated for the highest educational level in each round, is used.

The last dependent variable that is included is the effect of education on status, or the *returns of education*. Here the regression coefficient of achieved status on education is used, and again (because of the changing scale of education) this is the standardized coefficient. Summary statistics for the dependent variables are displayed in the table below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEO</td>
<td>0.262</td>
<td>0.107</td>
<td>0.010</td>
<td>0.535</td>
</tr>
<tr>
<td>Conditional primary</td>
<td>0.185</td>
<td>0.180</td>
<td>-0.219</td>
<td>0.554</td>
</tr>
<tr>
<td>ISO</td>
<td>0.187</td>
<td>0.107</td>
<td>-0.113</td>
<td>0.493</td>
</tr>
<tr>
<td>Returns of education</td>
<td>0.608</td>
<td>0.257</td>
<td>-0.598</td>
<td>0.971</td>
</tr>
</tbody>
</table>

### 4.2 Independent variables

Here again, there are several possibilities when it comes to choosing appropriate independent variables. A major distinction is between “real” parameter values and variables that are the result of these values.

In some cases, it is clear that the parameter values as they are used in the simulation can without any problems be used as independent variables. This is the case for the number of educational levels, the meritocratic parameter and the proportion of high status positions.

In other cases however this is less clear. For instance, the relation between social origin and learning skills (the primary effect) is governed by a parameter varying between .1 and .5. These different values result in a varying association between origin and achievement, that can be expressed in for example a regression coefficient. Now the question is whether the parameter value or the regression coefficient should be used as independent variables in the analyses.

One reason for the use of parameter values would be that they accurately reflect the mechanisms that produce the outcomes of the simulation process, as described in chapter three. On the other hand, they do not have a meaning that can easily be interpreted substantively; they are relatively arbitrary means of producing different associations between variables. Regression coefficients in contrast, being commonly used in empirical research, can be interpreted without much problems.

The decisive argument however for using regression coefficients rather then parameter values as independent variables is that the analysis of the simulation outcomes is ultimately meant to produce testable hypotheses. Regression coefficients (or other statistical measures) can be measured in actually existing cohorts; parameter values of course cannot.
Thus, for the primary and secondary effects of stratification the regression coefficients of achievement on origin and survival probabilities on origin (controlled for the effect of achievement) are used as independent variables (see figure 4.1).

In the case of the secondary effect, there is one complication that is related to the way in which educational expansion is implemented in the model. Because educational expansion is implemented as an overall increase of the survival probabilities, using a ceiling effect, this also directly influences the coefficient of the secondary effect. To be able to estimate the effect of expansion while controlling for the variation in the secondary effect, only the coefficient of the secondary effect from each first cohort should be considered. However, this method has its disadvantages too. The most problematic aspect of this variable is that it will be hard to measure it empirically; it is not clear how one should measure this effect as it would have been if no educational expansion had taken place. Therefore, the analysis will be conducted with both types of the secondary effect.

Like educational inequality, educational expansion is a concept that can be expressed in different ways. Rijken (1999) for example proposes to use both the mean and the standard deviation of education as measures of educational expansion. This measure would be problematic in the context of this study, because of the varying scale of education (discussed earlier). Moreover, as educational expansion is generally described as increasing enrollment rates, there is a better measure available in the simulation data: the average survival probability over all cases at each round.

For the analysis of ISO, there are in principle three independent variables: the meritocratic parameter, IEO, and the distribution of available status positions.

Of these three the first is most easily included in the analyses: one can simply use the parameter value. IEO has already been discussed as a dependent variable, and the same variable that was chosen as a dependent variable can now be used as an independent variable. For the distribution of status positions, a variable is produced by the simulation program expressing the proportion of high status positions. This variable will be used in the analyses.

All independent variables used in the analyses are summarized in table 4.2. In this table, the corresponding values that were used in Boudon’s original numerical example are also displayed. In most cases, the values used in the simulation are more or less equally distributed around the values used by Boudon, with the exception of the number of educational levels and the distribution of status positions. Boudon constructed his example such that the number of educational levels would be relatively high, and there are indications that this number is indeed higher than observed in actually existing educational systems (see for example Halsey et al. 1980). Therefore only lower numbers of educational levels are examined. With regard to the distribution of status positions, I start with Boudon’s original 1000-3000-6000 distribution, and then increase the proportion of high status positions to simulate a modernization process.
Figure 4.1: primary and secondary effects

![Diagram showing the relationship between Origin, Primary, Secondary, Achievement/skills, and Survival prob.]

Table 4.2: Descriptive statistics of Independent variables in the simulation, with the values used by Boudon

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Boudon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educ. levels</td>
<td>5.000</td>
<td>2.000</td>
<td>2.000</td>
<td>8.000</td>
<td>9</td>
</tr>
<tr>
<td>Educ. Levels Centered</td>
<td>0.000</td>
<td>2.000</td>
<td>-3.000</td>
<td>3.000</td>
<td>3</td>
</tr>
<tr>
<td>Primary</td>
<td>0.300</td>
<td>0.141</td>
<td>0.064</td>
<td>0.520</td>
<td>0.267</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.085</td>
<td>0.049</td>
<td>0.007</td>
<td>0.191</td>
<td>-</td>
</tr>
<tr>
<td>Secondary first cohort</td>
<td>0.130</td>
<td>0.046</td>
<td>0.054</td>
<td>0.191</td>
<td>0.157</td>
</tr>
<tr>
<td>Expansion</td>
<td>0.611</td>
<td>0.171</td>
<td>0.360</td>
<td>0.933</td>
<td>0.503</td>
</tr>
<tr>
<td>Expansion centered</td>
<td>0.211</td>
<td>0.171</td>
<td>-0.041</td>
<td>0.532</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>0.262</td>
<td>0.107</td>
<td>0.010</td>
<td>0.535</td>
<td>0.405</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.750</td>
<td>0.171</td>
<td>0.500</td>
<td>1.000</td>
<td>0.7</td>
</tr>
<tr>
<td>% High status positions</td>
<td>0.19</td>
<td>0.067</td>
<td>0.100</td>
<td>0.280</td>
<td>0.1</td>
</tr>
</tbody>
</table>

The values as used by Boudon are recalculated from Boudon’s simulated data for the first cohort.

4.3 Results: IEO

In the remainder of this chapter, the results of the simulation will be presented in the order of the research questions. The first set of results concerns IEO, and more specifically the following research questions:

**Research question 1:**

What is the influence of the association between parental status and achievement level on IEO?

**Research question 2:**

What is the influence of the (controlled) association between parental status and the survival probability on IEO?

**Research question 3:**

What is the influence of the number of educational levels on IEO?

**Research question 5:** To what extent do the effects of the number of educational levels, the primary and secondary effect, and educational expansion interact?
Table 4.3 presents the results of three regression models where the standardized regression coefficient of education on origin (for IEO) is the dependent variable (as discussed earlier) and the number of educational levels, the regression coefficients of survival probabilities (the secondary effect) and achievement (the primary effect) on origin, and educational expansion (the mean of the survival probabilities) are the independent variables. For convenience in the following analyses, the number of educational levels is centered around it’s mean (5). Where educational expansion is not considered yet (model 1 and 2), only observations from every first cohort are taken into account. Due to the high number of “observations”, all displayed coefficients are significant at the 5%- level.

Table 4.3: Regression models with IEO as dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>beta</td>
<td>Coeff.</td>
</tr>
<tr>
<td>Constant</td>
<td>0.050</td>
<td>-</td>
<td>-0.024</td>
</tr>
<tr>
<td>Nr. Educ levels</td>
<td>0.014</td>
<td>0.300</td>
<td>0.003</td>
</tr>
<tr>
<td>Primary</td>
<td>0.414</td>
<td>0.621</td>
<td>0.666</td>
</tr>
<tr>
<td>Secondary</td>
<td>1.315</td>
<td>0.637</td>
<td>1.884</td>
</tr>
<tr>
<td>Expansion</td>
<td>-</td>
<td>-</td>
<td>-1.30</td>
</tr>
<tr>
<td>Nr. Educ*Primary</td>
<td>0.014</td>
<td>0.098</td>
<td>-0.130</td>
</tr>
<tr>
<td>Nr. Educ*Secondary</td>
<td>0.054</td>
<td>0.157</td>
<td>0.024</td>
</tr>
<tr>
<td>Primary*Secondary</td>
<td>-1.916</td>
<td>-0.499</td>
<td>0.043</td>
</tr>
<tr>
<td>R²</td>
<td>0.93</td>
<td>0.96</td>
<td>0.94</td>
</tr>
<tr>
<td>N</td>
<td>21000</td>
<td>21000</td>
<td>84000</td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level.

The table shows that the number of educational levels, the primary effect and the secondary effect all have a positive effect on IEO (model 1). The exact interpretation of the absolute coefficients is however not that straightforward, because of the complex measures involved. The unit of the independent variable IEO, to start with, is the effect of a change of one standard deviation of social origin in terms of the standard deviation of education (if IEO is 0.5, an one more standard deviation of social origin results in half a standard of education more). An additional educational level thus has the effect that the effect of a change of one standard deviation of origin on education increases with 0.014 standard deviations of education. The primary effect is the regression coefficient of skills on social origin (both on a 1-3 scale); the coefficient of the primary effect in the table above means that if the effect of social origin on skills increases with one, IEO increases with 0.414. The secondary effect is expressed as the regression coefficient of the survival probability on social origin: thus if the direct effect of social origin on the survival probability increases with (say) 0.1, IEO increases with 0.13.

In terms of standardized coefficients, the secondary effect and primary effect are almost equally important, while the number of educational levels is less important. The interpretation of these standardized coefficients is somewhat problematic however. They are calculated using the ratio of the standard deviations of the variables involved, but these deviations result from distributions that are not observed in reality but are more or less arbitrarily chosen. Thus the standardized coefficients only have
a meaning that is related to these artificial distributions within the simulation, but none outside this context.

Note furthermore that the explained variance is very high: the adjusted $R^2$ is .93. This is due to the fact that these data are simulated data, where no unknown causes or unobserved variables exist: all the variables that can play a role in our simulated world are included in the analysis, we only have to find out how these variables have an effect.

It can be expected that except for the effect now described there are also important interaction effects at work, because in theory the primary and secondary effect only work via a number of educational levels. Therefore in model 2 interaction effects of all three variables are included.

The inclusion of interaction effects results in a large reduction of the value of the coefficient of the number of educational levels. This coefficient can now be interpreted as the effect of the number of levels when the primary and secondary effects are both equal to zero, and is as a result close to zero itself (however still significant, due to the large number of cases). The main effects of the primary and secondary effects can now be interpreted as the effects at the mean number of educational levels (5) and at each other’s zero value. Both these effects increase with an increasing number of levels: the (unstandardized) primary effect changes from a minimum of .62 to a maximum of .71 (given that the secondary effect is zero); the secondary effect changes from 1.72 to 2.05 (given that the primary effect is zero).\(^5\)

The negative interaction effect of the primary and secondary effect shows that these two effects have a diminishing effect on each other. At the mean number of educational levels, the primary effect varies between .30 and .67; the secondary effect varies between .89 and 1.75.

These analyses result in the following hypotheses:

\begin{itemize}
  \item \textit{Hypothesis 1: Given that there is an effect of social origin on survival probabilities, an increase in the number of branch points will result in an increase in IEO.}
  \item \textit{Hypothesis 2: Given that there are branch points, an increase of the direct effect of social origin on survival probabilities (secondary effect) results in an increase in IEO.}
  \item \textit{Hypothesis 3: Given that there are branch points, an increase of the effect of social origin on school achievement (primary effect) results in an increase in IEO.}
\end{itemize}

\(^5\) For example: when the number of transitions is at it’s (centered) minimum of -3, the size of the primary effect is $0.666 - 3*0.014 = 0.624$.  

35
Hypothesis 4: The larger the number of branch points, the larger the positive effect of the primary and secondary effects on IEO.

Hypothesis 5: The larger the primary effect, the smaller the secondary effect, and vise versa.

The meaning of hypotheses 1 to 4 is not hard to understand intuitively: since IEO is the result of an effect of social origin on the survival probabilities at each transition (through the primary and secondary effects), it is not very surprising that IEO increases when these mechanisms have more opportunities to have an effect. Conversely, an increase of the number of transitions will only matter if there is really something happening at these transitions with regard to IEO, i.e. if there are primary and secondary effects at work. The stronger these effects, the stronger the effect of an additional transition will be. The interaction between the primary and secondary effect is more surprising: it is hard to imagine why the primary and secondary effects would have an effect on each other. On the other end, the very purpose of a simulation model is to produce theoretical results that were not easily expected beforehand. I will pay more attention to this point in later chapters.

Educational expansion and selection

I now turn to the following research question:

Research question 4:

What is the influence of educational expansion on IEO?

The effect of educational expansion is first investigated by adding the mean of the survival probabilities as an independent variable to the model, and extending the analysis to all observations (instead of only the first cohorts). The results are described in table 3, model 3. They show in the first place that educational expansion – expressed as the average survival probability- has a negative effect on IEO: the higher the average survival probability, the lower IEO. The other effects in the model have generally become smaller and the interaction between the primary and secondary effect is now positive. However, as discussed above under “independent variables”, the effect of the secondary effect might be biased in this model because its measure is influenced by educational expansion. To avoid this problem, model 4 in table 4 shows a model where the secondary effect is included as the regression coefficient of the survival probabilities on social origin of the first cohort. This means that for every sequence of four cohorts, the secondary effect is assumed to be as it was in the first cohort, where it was not influenced by educational expansion. All the interaction effects refer to this variable (Secondary c1).
Table 4.4: Regression models with IEO as the dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>Coeff.</th>
<th>beta</th>
<th>Coeff.</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.279</td>
<td>0.025</td>
<td>0.001</td>
<td>0.019</td>
</tr>
<tr>
<td>Nr. Educ. levels</td>
<td>0.003</td>
<td>0.054</td>
<td>0.001</td>
<td>0.019</td>
</tr>
<tr>
<td>Primary</td>
<td>0.443</td>
<td>0.583</td>
<td>0.572</td>
<td>0.753</td>
</tr>
<tr>
<td>Secondary c1</td>
<td>1.234</td>
<td>0.525</td>
<td>1.631</td>
<td>0.694</td>
</tr>
<tr>
<td>Expansion</td>
<td>-0.427</td>
<td>-0.677</td>
<td>-0.089</td>
<td>-0.141</td>
</tr>
<tr>
<td>Nr. Educ*Primary</td>
<td>0.023</td>
<td>0.139</td>
<td>0.023</td>
<td>0.142</td>
</tr>
<tr>
<td>Nr. Educ * Sec. C1</td>
<td>-1.268</td>
<td>0.179</td>
<td>0.071</td>
<td>0.184</td>
</tr>
<tr>
<td>Primary * Sec. C1</td>
<td>-1.268</td>
<td>-0.290</td>
<td>-1.513</td>
<td>-0.346</td>
</tr>
<tr>
<td>Nr. Educ*Expansion</td>
<td>0.007</td>
<td>0.036</td>
<td>0.007</td>
<td>0.036</td>
</tr>
<tr>
<td>Primary*Expansion</td>
<td>-0.455</td>
<td>-0.290</td>
<td>-1.487</td>
<td>-0.351</td>
</tr>
<tr>
<td>Sec. C1*Expansion</td>
<td>-1.487</td>
<td>-0.290</td>
<td>-1.513</td>
<td>-0.346</td>
</tr>
</tbody>
</table>

R² 0.94 0.97
N 84000 84000

All displayed coefficients are significant at the 5% level.

The manipulation described above seems to affect the results to some extent. In the first place, the effect of educational expansion decreases, (both in absolute value and in terms of the standardized coefficient,) while the effects of the primary effect and expansion increase. Second, the interaction between the primary and secondary effects is negative again, as in model 2.

The analysis of the effect of educational expansion on IEO could be extended by adding interaction effects with educational expansion to the model. It would then be however be convenient to transform this variable as to avoid that the main effects of the other variables would refer to a situation in which the mean of the survival probabilities would be equal to zero. A reasonable way to do this is to centre the variable around its mean for all first cohorts, in which case the coefficients of the other variables refer to the situation in which educational expansion does not play any role (namely, in all fist cohorts). The results of this analysis make up model 5.

The three interaction effects show that there are is indeed interaction between educational expansion and the other three variables. The interaction effects with the primary and secondary effects are both negative, meaning that the (negative) effect of educational expansion on IEO becomes stronger, the larger the primary and secondary effects on school continuation. The effect of the number of educational levels on the effect of expansion is positive but close to zero, however still significant.

The fact that educational expansion as implemented in the simulation model has a negative effect on IEO appears to be in line with Mare’s hypothesis that educational expansion should have a negative effect on a linear regression coefficient. Also, it appears from the table above that the negative effect of an increase of the survival probabilities could be offset by an increase of the primary and secondary effects. For instance, an increase of the survival probabilities with one standard deviation would be offset by an increase of either the primary or the secondary effect with about one standard deviation. A
further relation between Mare’s and Boudon’s theories is investigated in the following research question:

Research question 7: What is the effect of educational expansion on the association between social origin and school achievement at the highest educational level?

To answer this question the regression coefficient of school achievement on social origin for the highest level (or conditional primary effect, as discussed above) is used as the dependent variable. The independent variables are in the first place educational expansion, with the primary effect and the number of educational levels as control variables.

Table 4.5: Regression model with the conditional primary effect at the highest educational level as dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>Coeff</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.303</td>
<td>-</td>
</tr>
<tr>
<td>Expansion</td>
<td>0.539</td>
<td>0.511</td>
</tr>
<tr>
<td>Primary</td>
<td>0.977</td>
<td>0.767</td>
</tr>
<tr>
<td>Nr. Educ. Levs</td>
<td>-0.027</td>
<td>-0.300</td>
</tr>
<tr>
<td>R²</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>84000</td>
<td></td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level

The effect of educational expansion appears to be positive, which is in line with Mare’s argument that the effect of origin on school achievement (skills) increases after expansion because variation on achievement increases. Thus we see that Boudon’s model of stratification seems to lead to the same conclusions as Mare’s model.

The analyses above can be summarized in the following hypotheses:

Hypothesis 6: The larger educational expansion, the smaller IEO.

Hypothesis 7: The larger the primary effect, the stronger the negative effect of educational expansion on IEO.

Hypothesis 8: The larger the secondary effect, the stronger the negative effect of educational expansion on IEO.

Hypothesis 9: The larger educational expansion, the stronger the effect of origin on achievement at the highest educational level.
4.4 Results: ISO

In the following the research questions concerning the inequality of status opportunity are considered. The questions, as formulated before, were:

*Research question 8: how does the level of ISO depend on the level of IEO?*

*Research question 9: how does the level of ISO depend the value of the meritocratic parameter?*

*Research question 10: What is the influence of change in the distribution of status positions on ISO?*

*Research question 11: To what extend do the effects of IEO, the meritocratic parameter and the status distribution interact?*

Model 1 in table 4.6 below results from a regression analysis using the (uncontrolled) regression coefficient of social origin on achieved status as the dependent variable and the IEO, the meritocratic parameter and the change in the status distribution as independent variables (see under “independent variables” above). Only the rounds where there was no problem in the distribution process (see “The simulation model”) are taken into account.

Table 4.6: Regression model with ISO as the dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>1 Coeff.</th>
<th>Beta</th>
<th>2 Coeff.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.311</td>
<td>-</td>
<td>-0.096</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>0.823</td>
<td>0.777</td>
<td>0.935</td>
<td>0.882</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.345</td>
<td>0.550</td>
<td>0.021</td>
<td>0.033</td>
</tr>
<tr>
<td>Prop. high status</td>
<td>0.078</td>
<td>0.049</td>
<td>0.291</td>
<td>0.183</td>
</tr>
<tr>
<td>IEO*Prop. high</td>
<td>-0.843</td>
<td>-</td>
<td>-0.220</td>
<td>-</td>
</tr>
<tr>
<td>IEO*Merit</td>
<td>0.788</td>
<td>0.379</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merit*Prop. high</td>
<td>0.527</td>
<td>0.170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.88</td>
<td>0.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>71668</td>
<td>71668</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level. In model 2, the level of meritocracy is centered at 0.7.

The results show in the first place that, unlike Boudon claimed, there is an effect of IEO on ISO. In terms of beta coefficients, this effect is even relatively large.

Second, in line with Müller-Benedicts findings, the meritocratic parameter appears to have an positive effect on ISO: the higher the meritocratic parameter, the higher status inequality. The size of the coefficient means that if meritocracy would increase from 0.5 (a high level of education only gives a 50% change on a high status job) to 1, the effect of social origin on status (both expressed on a 1-3 scale) increases with 0.173.
Third, and most surprisingly, the amount of change in the status distribution appears to have a positive effect on ISO. This is somewhat surprising because one would expect that a relatively larger number of high status positions would facilitate mobility. The explanation of this positive effect becomes clear when interaction effects are included in the model resulting in model 2.

To allow for an easier interpretation, the variable for meritocracy has been centered at .7, so that the main effects of the other variables can be interpreted as the effects when the value of the meritocratic parameter is .7, which is the value used by Boudon and is also close to the mean of the variable (.75). The table shows that there is a rather strong interaction effect between IEO and the proportion of high status positions (distribution effect). At a meritocratic level of .7, the effect of the proportion of high status positions is positive for low levels of IEO, but becomes negative when the IEO-variable becomes larger then .34. This relation however is also dependent on the level of the meritocratic parameter: the higher the level of meritocracy, the higher the “turning point” of the distribution effect to become negative. For highest level of meritocracy (where the value of the parameter is 1, representing perfect meritocracy), the distribution effect is even positive for all levels of IEO.

The explanation for these effects could be as follows. In situations with a high level of meritocracy, a relatively low level of IEO, and few high status positions, there are many candidates from a high background who cannot get a high status position because there is a lot of competition from the lower classes. Now when the number of available high status positions increases, the candidates with a high background have a higher chance of maintaining their status (being immobile) despite a relatively low level of education, and thus ISO increases.

In situations with a high level of IEO however, candidates with a high social background are already relatively sure of a high status position. An increase in the number of high status positions would then allow more low background candidates into high status positions, and ISO would decrease.

The positive interaction effect between meritocracy and changing of the distribution would reflect that a higher level of meritocracy causes a stronger competition, which increases the positive effect of the distribution as described above.

The finding that a change in the status distribution towards more high status jobs can, under certain circumstances, lead to a higher level of ISO, casts new light on the idea that the stability of status inequality predicted by Boudon is caused by the assumed stability of the status distribution. The results above show that, in theory, ISO could remain stable because the relative number of high status positions increases. Namely, when IEO decreases but at the same time (or perhaps even as a result of the change in IEO) the proportion of high status positions increases, these two trends could – under certain circumstances - offset each other’s effects, resulting in a stable level of ISO. A further decline of IEO would then lead to a stronger effect of the distribution change, etc. Thus the two processes of a decline in IEO and a change towards more high status positions would, according to this model, lead to a quite complex development of ISO.
To sum up, the above findings lead to the formulation of the following hypotheses:

*Hypothesis 10:* The higher IEO, the higher ISO.

*Hypothesis 11:* The higher the level of meritocracy, the higher ISO.

*Hypothesis 12:* The higher the level of meritocracy, the higher the positive effect of IEO on ISO.

*Hypothesis 13:* When the level of meritocracy is high and IEO is low, an increase in the proportion of high status positions has a positive effect on ISO. When meritocracy is low and IEO is high, the effect of an increase of the proportion of high status jobs is negative.

The returns of education

The last research question considered here is about the effect of education on status:

*Research question 12:* How is the effect of education on occupational status influenced by the level of meritocracy, IEO and the distribution of status positions?

This question is investigated by taking the (standardized) regression coefficient of status on education (or the returns of education) as a dependent variable, and taking the same variables as in the previous analyses as the independent variables. The results are displayed in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Coeff.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.105</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>1.680</td>
<td>0.658</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>1.063</td>
<td>0.704</td>
</tr>
<tr>
<td>Prop. High status</td>
<td>1.088</td>
<td>0.283</td>
</tr>
<tr>
<td>IEO*P. high</td>
<td>-4.437</td>
<td>-0.481</td>
</tr>
<tr>
<td>IEO*Merit</td>
<td>-0.601</td>
<td>-0.120</td>
</tr>
<tr>
<td>Merit*P. high</td>
<td>0.816</td>
<td>0.109</td>
</tr>
</tbody>
</table>

R² = 0.58
N = 71668

All displayed coefficients are significant at the 5% level

The results show that IEO, meritocracy as well as the proportion of high status positions positively influence the effect of education on status. There are however strong interaction effects: at high levels
of IEO (more than .38) the effect of the proportion of high status positions becomes negative. Also, the effect of IEO becomes much smaller with higher levels of meritocracy (and vice versa), while the effect of meritocracy becomes larger with a higher proportion of high status positions.

The fact that there is a strong positive effect of IEO on the effect of education on status indicates that there is something like devaluation of educational degrees in the model: as inequality of educational opportunity goes down, reaching a high status position becomes more difficult with a given level of education. In this sense Boudon’s model is in line with Thurow’s “labor queue” model (Thurow 1975, Wolbers 1998).

When we look at the effect of the proportion of high status positions, the pattern looks similar to the pattern we saw while analyzing the effect on ISO. Again the effect is positive for low levels of IEO and high levels of meritocracy, and negative for high levels of IEO and low levels of meritocracy.

**Hypothesis 14:** the weaker IEO, the weaker the effect of education on status.

**Hypothesis 15:** the stronger meritocracy, the stronger the effect of education on status.

**Hypothesis 16:** the higher IEO, the weaker the effect of meritocracy on the effect of education on status.

**Hypothesis 17:** When the level of meritocracy is high and IEO is low, an increase in the proportion of high status positions has a positive effect on the effect of education on status. When meritocracy is low and IEO is high, the effect of an increase of the proportion of high status jobs is negative.
4.5 Conclusions

In order to explore the properties and consequences of Boudon’s stratification theory, the data generated by the simulation program were analyzed in this chapter. The results of this analysis were formulated as hypotheses and can be regarded as theory development. Some of these hypotheses are quite straightforward and not very surprising, like the finding that stronger primary and secondary effects increase inequality, and that these effects become stronger when they are allowed to work at more transitions. Other findings appear to be in line with Mare’s theory: what is left of the primary effect at the end of the educational career becomes stronger when the educational system expands. Müller-Benedicts findings that ISO increases with more IEO and more meritocracy could also be confirmed by my analysis. The most interesting result is probably the changing effect of extra high status positions on ISO, because it results from an extension of the simulation model and is quite counterintuitive.

Although the hypotheses were stated verbally and qualitatively, note that these are actually just simplifying verbal expressions of the regression models, which are much more precise and might perhaps be considered as the “real” hypothesis. In the empirical part of this study, they will be used again as such to generate predictions using real data.
5. Testing the Model

In this chapter, the hypotheses that were generated by the simulation model in the previous chapter will be subjected to empirical testing. But first, a few problems connected to the testing of this model have to be discussed.

In the first place, there is the problem of a “time inconsistency” within the assumptions of the model. The theory pretends to give a description of the mechanisms underlying the complete stratification process from social origin, via education, to social status. If that were true, one could then test the theory by comparing the predictions to survey data on people who have completed the whole process. This is however prohibited by Boudon’s assumption that people who start their educational career at the same moment, will also enter the labor market at the same time. This is obviously an unrealistic assumption, because higher education just takes more time than lower education. Whatever the theoretical consequences of this assumption, it has at least the methodological consequence that it is not possible to submit both the IEO and ISO models to data on the same cohort, because the members of one educational cohort will not also be one labour market cohort. For this reason, the predictions on IEO and ISO will be tested separately, using different datasets.

A second problem is that it is sometimes difficult to test the hypotheses that were formulated in the previous chapter separately. The reason is that they are all verbal approximations of complex regression models. These models give very precise quantitative information on the expected outcomes, but the relations are sometimes so complex (due to the interaction effects) that it can be hard to say something about the effect of a single variable. The hypotheses on a certain outcome must be considered as a system (also because they are the products of one coherent system – the simulation model), and I hope the approach to the testing of the model explained below will solve that problem.

5.1 Testing IEO predictions

Data

In order to be able to test the predictions of the IEO model, we need data with at least the following information on individuals: social origin, learning abilities (to measure primary effects), and the course of the educational career. Unfortunately, while datasets which include information on social origins and educational attainment are widely available, few of these also incorporate information about learning skills, probably due to the intensive measurement methods needed. Among the exceptions are the four sets of Dutch educational panel data analyzed here. These are the “School Career and Background of Pupils in Secondary” (SMVO; CBS 1984) surveys of 1976 and 1982, and the “Cohort of Students in Secondary Education” surveys of 1989 and 1993 (collected by the Central Bureau of Statistics, available through the Scientific Statistical Agency). In these studies four cohorts
of students are followed during their educational career, starting from their first year of secondary education. In the SMVO data the students were followed for respectively thirteen and sixteen years, in the VOCL data for twelve and eight years. This means that in the 1993 cohort, it is only observed if people entered higher education, but not if they completed it. The four panel studies were combined into one dataset to be able to analyze them simultaneously.

Method
To find out whether the theoretical model is able to predict outcomes in reality, the outcomes of the simulation model can be compared to those of the SMVO/VOCL data. To achieve this we can treat the four cohorts as if they were simulated cohorts, and see what the simulation model would predict for cohorts with these characteristics, using the regression models developed in chapter four. With such a small number of observations (only four cohorts), it is not possible to test hypotheses by estimating the expected effects from the data.

To be able to compare the IEO patterns in these four cohorts to those that were observed in the simulation data it is necessary to construct dependent and independent variables that are as similar as possible to the measures that are applied in the analyses of the simulation data. To achieve this the scales of the variables of the survey data will be adapted to match those used in the simulation data.

Variables

Variables on the individual level
The independent variables on cohort level that are used in the simulation analyses are the strength of the primary effect, the strength of the secondary effect, educational expansion, and the number of educational levels. On the individual level, the variables that underlie these cohort level variables are social origin, skills, the survival probability, and educational attainment. In the SMVO/VOCL data social origin was measured on a five-point scale, where the lowest is “worker” and the highest is “high employee”. The skills of the pupils were measured using a standard test that is commonly used in Dutch primary schools to assess pupils’ capabilities at the end of primary school, called the “CITO test”. This test measures both mathematical and reading skills, and the final skills variable that will be used in the remainder is constructed as the mean of these two scores. In the case of 1977, reading and math skills were measured on a different scale than in the other three cohorts (1-45 instead of 1-20). For this cohort, the variables were rescaled to the 1-20 scale of the other cohorts.

Educational attainment is included in the data as a five-point variable, where the different levels that existed between 1977 and 1993 were standardized in terms of the educational levels that existed in 1993. Thus the lowest level is “MAVO” or “LBO” and the highest level is a university degree. As will
be explained below, this variable had to be recoded into a three-level variable to be able to use it in the analyses. The independent variables on the individual level are summarized by cohort in the tables below.

Transitions
Between 1977 and 1993, the Dutch educational system has undergone many changes, however the main structure of transitions has remained more or less the same (de Graaf & Ganzeboom 1993). Following de Graaf and Ganzeboom, I define three transitions within the Dutch educational system: from primary education to lower secondary education (“MAVO” or “LBO”), from lower to higher secondary education (“HAVO”, “VWO” or “MBO”), and from higher secondary education to tertiary education (“HBO”, “WO”). For some pupils, the transition from lower to higher secondary education does not really take place, because they move from primary to higher secondary education in one step. However they are still assumed to have survived two transitions. Although there are thus three transitions within the Dutch educational system, the data at hand only contain two transitions, because the pupils of the four cohorts have been followed from their entrance in high school. This means that all of them have already survived the first transition, and only the transitions from lower to higher secondary education and form secondary to tertiary education are left to study.

Construction of cohort level variables
To be able to test the predictions made by the simulation model with the survey data, it is necessary to create cohort level variables that are as comparable as possible to those used in the simulation. The cohort level dependent variables needed are a measure for IEO, and a measure for IEO at the highest educational level. As independent variables we need measures for the primary and secondary effects, the number of educational levels, and a measure for educational expansion. A measure for IEO is created using the same method as in the simulation: it is defined as the standardized linear regression coefficient of the number of survived transitions on social background. To make the variables comparable however it is necessary to rescale the variables on the individual level to the scales used in the simulation, so social background is reduced from five values to a variable that ranges from 1 to 3. The other dependent variable in the analysis of IEO, the association between origin and skills at the highest level (the conditional primary effect) can easily be reproduced in the real data. Like in the simulation data, I use the regression coefficient of skills on origin, computed for those who have reached the highest educational level (which is in this case the third level).

The primary effect can be calculated by using the CITO test score as a measure for learning skills. Again I rescale this variable to the three-level scale used in the simulation data. The primary effect is then operationalized as the linear regression coefficient of the CITO test score on social origin.
The secondary effect is harder to establish. In the summation data, this variable is operationalized as the linear regression coefficient of the survival probability on social background, controlled for skills. In reality however, there is no such a thing as a survival probability available: students either survive a transition or not. The solution to this problem is to treat each transition per student as a single observation, and adding a variable that indicates whether the transition was survived (1) or not (0). Thus for students who have survived both transitions we have two observations, one for each transition, and scoring “1” on both. Students who survive only the first transition but fail at the second have two observations as well, scoring “1” at the first and “0” at the second transition. Those who do not survive the first transition generate only one observation, scoring “0”. Ideally, one would at this point use logistic regression analysis to estimate the (controlled) effect of social background on surviving a transition, and use this estimate as a measure for the secondary effect. In the simulation data however the secondary effect is represented as the linear regression coefficient of the survival probability, and in order to approximate this measure as good as possible I will use linear regression in the survey data as well. Thus the secondary effect is operationalized as the linear regression coefficient of surviving a transition on social background, controlled for skills.

The survival probabilities are in the simulation data also used to construct the measure for educational expansion: the “size” of the educational system is described as the mean of the individual survival probabilities. To reproduce this measure in the four observed cohorts, we can again use the set of transitions that was used to calculate the secondary effect: educational expansion is then the proportion of successfully completed transitions within each cohort.

The number of educational levels, to conclude the discussion of independent variables, is constant in all four cohorts and equals three. However, in order to make the analysis of these data comparable to the analysis of the simulation data, the number of educational levels should be centralized around the mean of educational levels in the simulation data, which is five. The centralized number of educational levels in the SMVO/VOCL data is then –2. (Which means that the number of levels is relatively small in terms of the model.)
**Results: primary and secondary effects; expansion**

The cohort level variables that were constructed using the methods above are displayed in the table below.

**Table 5.1: Dependent and independent variables in four Dutch education cohorts, 1977-1993**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cohort 1977</th>
<th>Cohort 1982</th>
<th>Cohort 1989</th>
<th>Cohort 1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEO</td>
<td>0.286</td>
<td>0.269</td>
<td>0.269</td>
<td>0.263</td>
</tr>
<tr>
<td>Conditional primary effect</td>
<td>0.027</td>
<td>0.035</td>
<td>0.040</td>
<td>0.046</td>
</tr>
<tr>
<td>Primary effect</td>
<td>0.133</td>
<td>0.120</td>
<td>0.135</td>
<td>0.134</td>
</tr>
<tr>
<td>Secondary effect</td>
<td>0.07</td>
<td>0.067</td>
<td>0.063</td>
<td>0.058</td>
</tr>
<tr>
<td>Expansion</td>
<td>0.514</td>
<td>0.567</td>
<td>0.564</td>
<td>0.534</td>
</tr>
</tbody>
</table>

The table shows in the first place that there are no strong differences between the four cohorts. Especially IEO seems to be more or less stable, with a small decrease between 1977 and 1989 and a small decrease again in the last cohort. The primary effect shows a downward trend in the first two cohorts, then increases in the 1989 cohort and remains stable. The secondary effect shows a slight decrease over all cohorts. The educational system seems to have expanded between 1977 and 1982, but then remained almost stable; it even shrunk somewhat between 1989 and 1993.

To be able to compare these findings to those from the simulation, we can apply the regression model that was estimated from the simulation data to the values of the independent variables of these four cohorts. The predicted IEO levels from this model can then be compared to those that were actually observed. The regression model used for this purpose is the full model (model 5) from chapter four as displayed in table 4.4 below.

**Table 4.4: Regression models with IEO as the dependent variable (simulated data)**

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>beta</td>
</tr>
<tr>
<td>Constant</td>
<td>0.279</td>
<td>0.025</td>
</tr>
<tr>
<td>Nr. Educ. levels</td>
<td>0.003</td>
<td>0.054</td>
</tr>
<tr>
<td>Primary</td>
<td>0.443</td>
<td>0.583</td>
</tr>
<tr>
<td>Secondary c1</td>
<td>1.234</td>
<td>0.525</td>
</tr>
<tr>
<td>Expansion</td>
<td>-0.427</td>
<td>-0.677</td>
</tr>
<tr>
<td>Nr. Educ*Primary</td>
<td>0.023</td>
<td>0.139</td>
</tr>
<tr>
<td>Nr. Educ * Sec. C1</td>
<td>-1.268</td>
<td>0.179</td>
</tr>
<tr>
<td>Primary * Sec. C1</td>
<td>-1.268</td>
<td>-0.290</td>
</tr>
<tr>
<td>Nr. Educ*Expansion</td>
<td>0.007</td>
<td>0.036</td>
</tr>
<tr>
<td>Primary*Expansion</td>
<td>-0.455</td>
<td>-0.268</td>
</tr>
<tr>
<td>Sec. C1*Expansion</td>
<td>-1.487</td>
<td>-0.351</td>
</tr>
<tr>
<td>R²</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>N</td>
<td>84000</td>
<td>84000</td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level.
The results from this comparison are displayed in table 5.3 and figure 5.1. In the first place, we see that the simulation model seems to underestimate the level of IEO in all cases. In the second place, the two lines shown in the figure seem to be more or less similar: first there is a decline (the size of which is exactly predicted by the simulation model), after which the level of IEO stays more or less stable. There are however some differences in the trend: the simulation model predicts a (very) small decrease over the last three cohorts, while in reality the level of IEO remained stable between 1982 and 1989 and declined thereafter. The main reason for the relatively sharp predicted decline between the first two cohorts seems to be the educational expansion between these two years: although the primary and secondary effects also decrease and therefore contribute to the decline of IEO, these effects are very small and would hardly cause any change in IEO just by themselves.

Table 5.2  Observed and predicted IEO, The Netherlands 1977 – 1993

<table>
<thead>
<tr>
<th>Cohort</th>
<th>1977</th>
<th>1982</th>
<th>1989</th>
<th>1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed IEO</td>
<td>0.286</td>
<td>0.269</td>
<td>0.269</td>
<td>0.263</td>
</tr>
<tr>
<td>Predicted IEO</td>
<td>0.191</td>
<td>0.168</td>
<td>0.170</td>
<td>0.170</td>
</tr>
<tr>
<td>Difference</td>
<td>0.098</td>
<td>0.101</td>
<td>0.099</td>
<td>0.092</td>
</tr>
</tbody>
</table>

Figure 5.1: Observed versus Predicted IEO, The Netherlands 1977-1993
Results: expansion & selection

In order to test the hypotheses concerning the relation between expansion and selection, the same method as applied above can be followed. This time the predicted association between origin and skills at the highest educational level will be compared to the association that is observed in the data. Like in the simulation analysis, the dependent variable is the regression coefficient of skills on status, for those who have reached the highest educational level (level three). The regression model that was used to predict the scores for the four cohorts is displayed below.

Table 4.5: Regression model with the conditional primary effect at the highest educational level as dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>Coeff.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.303</td>
<td>-</td>
</tr>
<tr>
<td>Expansion</td>
<td>0.539</td>
<td>0.511</td>
</tr>
<tr>
<td>Primary</td>
<td>0.977</td>
<td>0.767</td>
</tr>
<tr>
<td>Nr. Educ. Levs</td>
<td>-0.027</td>
<td>-0.300</td>
</tr>
<tr>
<td>R²</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>84000</td>
<td></td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level

The resulting predicted effects and the observed effects are shown in table 5.4 and figure 5.2 below. The trend in the observed effect of origin on skills on the highest level is clearly upwards, and results in an almost straight line. The predicted effects are also increasing, but sharper and only until 1989, after which there is a decrease.

The increase in the association between skills and origin as predicted by the model is not the result of an increase of the general primary effect: as is clear from table 5.1 above, the primary effect itself does not show a clear trend. That the effect of origin on skills at the highest educational level is nevertheless predicted to increase over the first three cohorts, is due to the interplay of the primary effect and expansion. Between the first and the second cohort, there is educational expansion, which is apparently strong enough to force the effect at the highest level upwards, although the primary effect itself is decreasing. From the second to the third cohort, there is no further expansion, but now the primary effect itself increases, which again causes a higher predicted primary effect at the highest level. From the third to the last cohort, there is negative expansion and a stable primary effect, which is predicted to lead to a decline of the dependent variable.

It is interesting that this upward trend, caused by the change in two different variables, is -at least qualitatively- confirmed by the data. The model however goes wrong about the 1993 cohort, where it predicts a decrease while in reality the upward trend is continued.
This concludes the testing of IEO predictions. Some of the hypotheses have not been tested directly: because the available data did not include any variation in the number of transitions, hypotheses related to the number of transitions could not be tested, including hypotheses on interaction effects with this variable. Hypotheses on the effects of the other variables could only be tested with a small number of observations. Support for the predictions is mixed: although the general predicted and observed patterns of the development of the dependent variables seem to be more or less alike, there are also differences. In the case of the association between origin and skills at the highest level, it is interesting to see that the observed increase can be explained by the changes in two different variables. The low number of observation (effectively four) makes it however difficult to make any judgements on the validity of the predictions. Fortunately, this will not be the case with the predictions concerned with ISO, which are the topic of the remainder of this chapter.

Figure 5.2: Observed and predicted conditional primary effect, The Netherlands 1977 – 1993
5.3 Testing ISO predictions

While testing the hypotheses on IEO was perhaps somewhat difficult because of the (un)availability of data with the necessary variables, this is not the case with ISO: large scale data with sufficient information were at hand. The testing of hypotheses on ISO will generally follow the same lines as the analysis of IEO described above (comparison of predicted and observed cohort level outcomes), but also take advantage of the larger number of observations by using a different method of testing the hypotheses: direct estimation of the hypothesized effects from the data.

Data

In the following analyses I will make use of (selections of) the so-called “International Social Mobility and Politics File” (ISMPF; Nieuwbeerta and Ganzeboom 1996) to test the hypotheses on ISO that were developed in the previous chapter. This dataset contains the combined information on stratification and mobility from a large number of surveys from 16 different countries over a long period of time (1956 – 1991), all made comparable. Among the variables included are social origin of respondents, their level of education, and the status of their current job. This, in combination with the large number of cases, the long time span and the international aspect makes this dataset an excellent means of comparing the outcomes of the model with the real world. In the following, I will start with the analyses of data on only one country (the Netherlands), and later extend the analysis to include more countries.

In analogy with the simulation analyses, the data will be analyzed by studying ISO outcomes for labor market cohorts. Because the data do not contain direct information however on labor market entry, this information has to be reconstructed. In order to do that I assume that all respondents started school at the age of four, and took their first job right after they left school. The year of labor market entry can then be reconstructed by adding four years and the number of years the respondent spent in school to his year of birth. This way, for each year a labor market cohort can be established of people who can be expected to have entered the labor market that year. In order to include a sufficient number of cases in each cohort I will limit the time period from 1945 (the end of the Second World War) to 1980.

Variables

For the beginning, the method applied here will initially be the same as for the hypothesis on IEO. For each labor cohort, outcome variables will be predicted by applying the regression models that resulted from the simulation analyses to independent variables that are taken from the data. These predicted outcomes can then be compared to real outcomes that are estimated for each cohort.

Like with the analysis of IEO above, it is again important to create variables that are in their scales comparable to the variables used in the simulation as much as possible. This is especially important because in the hypotheses on ISO interaction effects play an important role. For instance, it was hypothesised that an increase of the proportion of high status positions has different effects on ISO for
different levels of IEO. But if we want to apply the same regression models that resulted from the simulation to these real data, then it must be clear what level of IEO or meritocracy is high or low in relation to this regression model. In the following an attempt will be made to construct such variables.

Dependent variables

The first and most important dependent variable in this analysis is the size of ISO. In the simulation, ISO was defined as the linear regression coefficient of achieved status on social origin. The data contain information on the status of the respondent’s current or last job and on the status of his father’s job during adolescence, both measured at the International Social Economic Index (ISEI) (Ganzeboom et al., 1992). Because the theoretical model that underlies the simulation ignores career effects, we are basically only interested in the respondent’s first job, that he or she got right after he or she left the educational system. This information is - although it was included in the data- unfortunately missing for most cases. To somehow correct for career effects, I use age as a control variable, assuming that one’s status usually increases with age after one’s first job. ISO on cohort level is then the linear regression coefficient of the respondent’s status on father’s status, with age as a control variable. This results in a ISO variable which has more or less the same range as the corresponding variable in the simulation analyses.

The second dependent variable concerns the returns of education; the size of the influence of one’s level of education on one’s status. In the simulation analysis, this is the standardized regression coefficient of status on education. This measure can be reproduced from ISMPF-data by using the standardised regression coefficient of status (on the ISEI scale) on years of education, again while keeping age constant to reduce career effects. The result is a variable that is fairly comparable in range to its simulated equivalent (see tables 5.5 and 5.6).

Independent variables

In the analyses of the simulation results concerning ISO, I used three independent variables: IEO, meritocracy and the proportion of available high status positions.

Of these, IEO is the easiest to obtain from the data. In the simulation, IEO was the standardized coefficient of level of education on social origin. The level of education is included in the data as the number of years spent in school, which was constructed from all the different levels of education that were obtained by the respondents. For social origin, I use again the ISEI score of the father. IEO in this analysis is then the standardized linear regression coefficient of years of education on father’s status.

While IEO is relatively easy to construct from the data, meritocracy is much more problematic. The main reason for this is that meritocracy is actually a theoretical concept, referring to the value the level of education has for an employer who has to choose between applicants. In practise the actual value of meritocracy is hard to measure, however not impossible. Wolbers made a successful attempt to
estimate employers’ preferences for education by using log-linear models (on the same data), and found that the level of unemployment in a given period was an important predictor for this preference. In times when the labour supply is plenty (unemployment is high), employers apply a stricter selection on education then when labour is scarce. This explanation appeared to be more important then the competing explanation by modernization processes (Wolbers 1998). Therefore I will use the level of unemployment during the year of entry of each labour cohort as a proxy for meritocracy. Descriptive statistics on unemployment are displayed in table 5.5. The scaling problem however remains: the percentages of unemployment are much lower then the levels of meritocracy used in the simulation. This problem can in a very arbitrary manner, be solved by adding 0.5 to the levels of unemployment, which results in a variable for meritocracy with a range that more less resembles the range of the simulation variable (see tables 5.5 and 5.6). Note however that the exact values of this variable do not have any substantive interpretation; there is no reason why an unemployment rate of 20% should indicate a meritocratic value (as used in the simulation model) of 70%. Therefore, I will also conduct the analyses with other levels of meritocracy.

The last independent variable is the proportion of high status positions. In the simulation, this variable was simply the percentage of available status positions on the highest level of three, varying between 10% and 30%. In reality, it is not possible to determine exactly the number of available positions for a given labor cohort; only the distribution of job that were actually filled is observed. Thus I will use this distribution (that is, the distribution of respondent’s ISEI scores) to determine the share of high status positions in each cohort. The definition of “high” is of course somewhat arbitrary. In accordance with the simulation, I will define the upper third of the ISEI-range (10-90), which is everything above 63, as high. The resulting variable appears to resemble the simulation variable (see tables 5.5, 5.6).

Table 5.4: Cohort level dependent and independent variables in the Netherlands, 1946 – 1980

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO</td>
<td>0.299</td>
<td>0.066</td>
<td>0.008</td>
<td>0.424</td>
</tr>
<tr>
<td>Returns of education</td>
<td>0.580</td>
<td>0.045</td>
<td>0.482</td>
<td>0.665</td>
</tr>
<tr>
<td>IEO</td>
<td>0.300</td>
<td>0.073</td>
<td>0.095</td>
<td>0.421</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>8.196</td>
<td>7.663</td>
<td>0.900</td>
<td>28.200</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.582</td>
<td>0.077</td>
<td>0.509</td>
<td>0.782</td>
</tr>
<tr>
<td>Prop. High status</td>
<td>0.177</td>
<td>0.025</td>
<td>0.132</td>
<td>0.239</td>
</tr>
</tbody>
</table>

Table 5.5: Descriptive statistics of dependent and independent variables in the simulation analysis of ISO

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO</td>
<td>0.187</td>
<td>0.107</td>
<td>-0.113</td>
<td>0.493</td>
</tr>
<tr>
<td>Returns Educ.</td>
<td>0.608</td>
<td>0.257</td>
<td>-0.598</td>
<td>0.971</td>
</tr>
<tr>
<td>IEO</td>
<td>0.262</td>
<td>0.107</td>
<td>0.010</td>
<td>0.535</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.750</td>
<td>0.171</td>
<td>0.500</td>
<td>1.000</td>
</tr>
<tr>
<td>Prop. high status</td>
<td>0.150</td>
<td>0.112</td>
<td>0.000</td>
<td>0.300</td>
</tr>
</tbody>
</table>
Results: ISO

I will now describe the results of a comparison of the predictions made by the simulation model and the real outcomes on ISO in the Netherlands. To start with, the graphs in figures 5.3-4 below show the trends in the three independent variables: IEO, meritocracy and the proportion of high status positions. IEO shows an overall decline, especially in cohorts after the sixties. Meritocracy (which is actually just the unemployment rate plus 0.5) has first risen after the end of WWII until the beginning of the fifties, after which it has decreased until the sixties. Since then there has been a small increase. The percentage of high status positions show a mixed but overall upward trend.

Using the values of these three independent variables, we can now try to predict the level of ISO for each labor market cohort. Like before with IEO, the regression model developed in chapter four is used to generate these predictions. This is the full model including interaction effects, which is reproduced in the table below.

Table 4.6: Regression model with ISO as the dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>Coeff. 1</th>
<th>Beta 1</th>
<th>Coeff. 2</th>
<th>Beta 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.311</td>
<td>-</td>
<td>-0.096</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>0.823</td>
<td>0.777</td>
<td>0.935</td>
<td>0.882</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.345</td>
<td>0.550</td>
<td>0.021</td>
<td>0.033</td>
</tr>
<tr>
<td>Prop. high status</td>
<td>0.078</td>
<td>0.049</td>
<td>0.291</td>
<td>0.183</td>
</tr>
<tr>
<td>IEO*Prop. high</td>
<td></td>
<td></td>
<td>-0.843</td>
<td>-0.220</td>
</tr>
<tr>
<td>IEO*Merit</td>
<td>0.788</td>
<td></td>
<td>0.788</td>
<td>0.379</td>
</tr>
<tr>
<td>Merit*Prop. high</td>
<td></td>
<td></td>
<td>0.527</td>
<td>0.170</td>
</tr>
</tbody>
</table>

R² 0.88 0.90
N 71668 71668

All displayed coefficients are significant at the 5% level. In model 2, the level of meritocracy is centered at 0.7.

When applied to the values of the independent variables, this model predicts levels of ISO for each cohort, that are plotted in figure 5.5 together with the observed values of inequality, with different ways to construct the variable for meritocracy. In figure 5.5, the line with ball symbols represents the observed level of ISO, which shows a clear overall downward trend. The line with triangle symbols shows the trend that is predicted when the variable on meritocracy as described above (unemployment + 0.5) is used. In the regression model with interactions, this variable was centered at the value of 0.7.

When using this variable, the level of ISO is clearly underestimated in nearly all cases. Because the exact relation between unemployment and meritocracy is unknown however, it is worth trying to use a higher level of meritocracy, in an attempt to move the predicted line closer to the observed line. On the other hand, the meritocracy scale cannot be shifted upwards too far, because that would result in levels of meritocracy that are more than 100%. Since the maximum unemployment rate observed in the data is about 28%, the maximum we can add to create a level of meritocracy is about 0.7 (0.7+0.28 = 0.98).

The predictions that result when meritocracy is calculated as unemployment + 0.7 are represented in the graph by the line with square symbols. This line comes indeed closer to the observed line, but still
ISO is generally underestimated. Note however that the underestimation of ISO seems to be more or less constant over time.

When we look at the shapes of the observed and predicted trends it appears that the two lines are in many regards remarkably similar. To begin with, both the predicted trend and the actually observed trend in ISO are on average decreasing from 1945 to 1980. But the observed trend also seems to follow the predicted trend on a more detailed level. This is especially clear in the first period from 1945 to 1960, where the shapes of the two lines are very similar. Also in the period from 1973 to 1980, the trend is predicted quite well, and now the predicted values are also somewhat closer to the observed values. During the period in between (1960-1970), the predictions are less accurate. The model here predicts a shaky increase and decline, with a maximum at the 1964 cohort. The observed trend on the other hand is more or less stable until 1965, and then declines. Another notable exception is the apparent drop of observed ISO in 1979, which is not predicted. The reason for this might be that the value of observed ISO for this cohort was not significant, maybe due to a relatively small number of observations within this cohort (143).

While the line with “high” meritocracy scale lies closer to the observed pattern, the correlation between the predicted and observed values is also a little higher: 0.72 for the “high scale line”, versus 0.69 for the “low scale line”.

The fact that the simulation model is able to predict the course of ISO development in the Netherlands reasonably well suggests that the mechanisms which underlie this model and which were identified in the hypotheses are indeed at work. We can try to explain the trend in ISO by taking a closer look at the graphs. The main factor in the generation of ISO seems to be IEO: there is a positive relation between IEO and ISO, in accordance with hypothesis 12. For instance, both the peaks in ISO in 1951 and 1956, as well as the minimum in the 1972 cohort, appear to more or less coincide with peaks respectively the minimum in IEO. In the period between 1960 and 1970 however this relation seems to be different. While the model continues to predict peaks in ISO in accordance with the pattern in IEO, the observed level in ISO remains rather stable. A possible explanation for this changing effect of IEO might be the dramatic drop in meritocracy during the sixties. According to hypothesis 14, the effect of IEO on ISO should be smaller the lower the level of meritocracy. Although the regression model that generated the predictions is supposed to take this effect into account, it is possible that the indicator for meritocracy (unemployment + 0.5) does not accurately reflect the real level of meritocracy. If in reality the difference in meritocracy between the 50’s and the 60’s was relatively larger then the values of the indicator suggest, then this might explain the smaller effect of IEO during the 60’s.

While the influence of IEO can be seen relatively easily in the graphs because it has such a strong influence, it is much harder to recognize the precise effects of the other two independent variables from the graphs, apart from the fact that in general the predictions – which are based on the combined hypotheses – seem to be in line with reality. Due to the complex interaction relations between the
variables that are incorporated in the predicting model, it is difficult to see what the precise effects of meritocracy and the proportion of high status positions are.

Another interesting observation that can be made from the graph is that the predicted development of ISO seems to precede the observed trend in many cases, for example in 1949 and 1955. This might indicate that in reality, a change in one of the independent variables needs more time to cause a change in ISO then it does in the simulation.

Figure 5.3: proportion of high status positions (scoring higher than 63 on the ISEI scale), The Netherlands 1946 -1980

Figure 5.4: level of meritocracy, calculated as unemployment rate + .5, The Netherlands 1946 -1980
Figure 5.5: IEO, The Netherlands 1945 – 1980

![Graph showing IEO values for different cohorts from 1945 to 1980.]

Figure 5.6: Observed and predicted ISO using different scales for meritocracy, The Netherlands 1946 - 1980

![Graph showing observed and predicted ISO values for different cohorts from 1945 to 1980, with different symbols and lines for observed and predicted values.]

58
Results: Returns of education

The second set of analyses related to the ISO-model concerns changes in the effect of education on achieved status, or the returns on education. The dependent variable for this analysis is the standardized regression coefficient of occupational status on education. As before, I will use the regression model developed in the simulation analysis to generate predictions on the development of this variable in the Netherlands during the period from 1945 to 1980. This regression model is reproduced by the table below.

Table 4.7: Regression model with the returns of education as the dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>Coeff.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.105</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>1.680</td>
<td>0.658</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>1.063</td>
<td>0.704</td>
</tr>
<tr>
<td>Prop. High status</td>
<td>1.088</td>
<td>0.283</td>
</tr>
<tr>
<td>IEO*P. high</td>
<td>-4.437</td>
<td>-0.481</td>
</tr>
<tr>
<td>IEO*Merit</td>
<td>-0.601</td>
<td>-0.120</td>
</tr>
<tr>
<td>Merit*P. high</td>
<td>0.816</td>
<td>0.109</td>
</tr>
<tr>
<td>R²</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>71668</td>
<td></td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level

The scaling problem of the meritocracy variable will again be dealt with by using different scales. I will use the same two versions of the meritocracy variable: unemployment + 0.5, and unemployment + 0.7, both centered at 0.7.

Figure 5.7 shows the observed and predicted trends for the returns of education. The line with ball symbols represents the observed pattern. The overall pattern appears to be almost horizontal, with a very minor decline over time. The value in 1980 is exceptionally low, but is significant. The line with triangle symbols represents the predictions that result when meritocracy is computed as unemployment + 0.5, and then centered at 0.7 like in the simulation analysis. These predictions appear to be very accurate until 1957, but after that the predicted trend shows an overall decline that is not observed in the real outcomes. Apart from that however, the shapes of the predicted trend and the observed trend are quite similar. The line with squared symbols represents the predictions that result when meritocracy is calculated as unemployment + 0.7 (and again centered). The shape of the resulting pattern is quite the same as before, but the line is shifted upwards. The returns of education are now overestimated until 1966, but are quite accurate after that.

The fact that the simulation model predicts a clear declining trend is the result of the strong decline of meritocracy (as indicated by unemployment) since the 1950’s, with meritocracy being the most powerful predictor in the simulation model. The increasing proportion of high status position even increases this effect of meritocracy. Apparently however, the influence of meritocracy has not been that strong in reality. As to the other variables, it is again hard to determine their precise effects on the predicted trend because of the interactions between them, and it is therefore not possible to test the
hypotheses on the development of the returns of education separately. They can however be tested as a system (described by the regression model), and as a system they seem to perform fairly well.

Figure 5.7: observed and predicted returns of education, using different scales of meritocracy, The Netherlands 1946 -1980

A different way of testing: estimation of effects

There is however a perhaps more precise way of testing the different hypotheses: one could try to estimate the effects found in the simulation analysis from the data, and thus reproduce the regression model that resulted from the simulation. This method is very common in the social sciences, but was not used for the analyses in IEO in this chapter because of the low number of cases: it is not possible to estimate a regression model on only four cases. The ISMPF – data however contain data on 13 countries over many years, and therefore a sufficiently large set of observations might be created. The data on only the Netherlands that were used to compare predictions and outcomes above still contain only 36 cases, and therefore data on four other countries are added: the United States, The United Kingdom, Germany and Australia. (For these countries, the number of cases per cohort was large enough to estimate IEO and ISO coefficients). Unemployment statistics on these countries of the years 1960, 1968 and 1971-1980 (OECD 1982) were used to compute meritocracy levels, resulting in 12 observations (labor market cohorts) for each of the four countries. Together with the data on the Netherlands, the resulting dataset contains 82 observations of labor market cohorts from five countries, with information on IEO, ISO, the level of meritocracy and the relative number of high status positions. IEO, ISO and the proportion of high status positions are calculated in the exact same

---

6 For other years between 1960 and 1971 no standardized unemployment statistics were available.
way as described in the previous section for the Netherlands. For meritocracy, the version that is
calculated as unemployment + 0.5 is used.

Table 5.6: Descriptive statistics on dependent and independent variables in five countries, 1946-1980

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO</td>
<td>82</td>
<td>0.249</td>
<td>0.083</td>
<td>0.424</td>
<td></td>
</tr>
<tr>
<td>IEO</td>
<td>82</td>
<td>0.299</td>
<td>0.072</td>
<td>0.453</td>
<td></td>
</tr>
<tr>
<td>Meritocracy</td>
<td>82</td>
<td>0.558</td>
<td>0.057</td>
<td>0.782</td>
<td></td>
</tr>
<tr>
<td>High status</td>
<td>82</td>
<td>0.173</td>
<td>0.053</td>
<td>0.250</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.8 presents the results of two regression analyses with ISO as the dependent variable and IEO, meritocracy and the proportion of high status positions as the independent variables. In model one, only main effects are included. All the effects are in the expected direction (see table 4.6) and significant at the 0.1 level (single tailed). The order of the size of the coefficients in terms of beta coefficients is also as predicted. These findings thus confirm hypotheses 10 & 11 on the effects of IEO and meritocracy (there was no hypothesis on only the main effect of the proportion of high status positions).

In model two also interaction effects are included. It appears however that the number of observations is still too small to estimate interaction effects: all p-values are higher then 0.6. Entering the interaction one by one into the model does not solve this problem. Therefore, the hypotheses on interaction effects can not be tested separately in this way.

For the second dependent variable in the ISO analyses, the returns of education, only the full model with interactions was estimated in the simulation analyses. For that reason I will refrain from repeating the analysis described in this section for this dependent variable.

Table 4.6: Regression model with ISO as the dependent variable (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>1 Coeff.</th>
<th>Beta</th>
<th>2 Coeff.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.311</td>
<td>-</td>
<td>-0.096</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>0.823</td>
<td>0.777</td>
<td>0.935</td>
<td>0.882</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.345</td>
<td>0.550</td>
<td>0.021</td>
<td>0.033</td>
</tr>
<tr>
<td>Prop. high status</td>
<td>0.078</td>
<td>0.049</td>
<td>0.291</td>
<td>0.183</td>
</tr>
<tr>
<td>IEO*Prop. high</td>
<td></td>
<td></td>
<td>-0.843</td>
<td>-0.220</td>
</tr>
<tr>
<td>IEO*Merit</td>
<td></td>
<td></td>
<td>0.788</td>
<td>0.379</td>
</tr>
<tr>
<td>Merit*Prop. high</td>
<td></td>
<td></td>
<td>0.527</td>
<td>0.170</td>
</tr>
<tr>
<td>R²</td>
<td>0.88</td>
<td></td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>71668</td>
<td></td>
<td>71668</td>
<td></td>
</tr>
</tbody>
</table>

All displayed coefficients are significant at the 5% level. In model 2, the level of meritocracy is centered at 0.7.
Table 5.7: regression models with ISO as the dependent variable in five countries, 1946-1980

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>Beta</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.140</td>
<td>-</td>
</tr>
<tr>
<td>IEO</td>
<td>0.357**</td>
<td>0.314</td>
</tr>
<tr>
<td>Meritocracy</td>
<td>0.407**</td>
<td>0.284</td>
</tr>
<tr>
<td>% high status</td>
<td>0.316*</td>
<td>0.155</td>
</tr>
<tr>
<td>IEO*% High</td>
<td>1.561</td>
<td>0.318</td>
</tr>
<tr>
<td>IEO*Merit</td>
<td>0.601</td>
<td>0.089</td>
</tr>
<tr>
<td>Merit*% High</td>
<td>0.601</td>
<td>0.089</td>
</tr>
<tr>
<td>R²</td>
<td>0.20</td>
<td>0.18</td>
</tr>
<tr>
<td>N</td>
<td>82</td>
<td>82</td>
</tr>
</tbody>
</table>

In model 2, the level of meritocracy is centered at 0.7
* = significant at the 0.1 level, ** = significant at the 0.05 level

5.3 Summary

In this chapter, Boudon’s models were tested in two different ways. In the first place, the predictions on IEO and ISO development that the models based on different independent variables observed in the Netherlands were compared to the observed trends. For IEO, the number of observations was very low, with makes it difficult to say much about the value of the hypotheses. However, the predicted and observed trends seemed to be roughly similar. For ISO, the number of observations was much larger, and allowed for a better comparison between observed and predicted trends. The model appeared to be able to predict trends in ISO and the returns of education quite well. However, because the predictions are the result of a system of hypotheses, it is still difficult to test the hypotheses separately. In the concluding analysis of this chapter, an attempt was made to test the hypotheses on ISO separately by directly estimating the expected effects from the data, after enlarging the number of observations by adding some more countries. Two hypotheses could be confirmed this way; testing further (interaction-) hypotheses on ISO would require even more observations. In the next and concluding chapter, I will discuss the consequences of the results for the validity of Boudon’s stratification models in more detail.
6. Conclusions & Discussion

6.1 Education, opportunity and social inequality

In the introduction of this study I started by asking the question whether a theoretical model could be found that explains, rather than just describes, the role of education in status attainment processes. The reason for asking this seemingly unoriginal question is that I had the impression that the majority of stratification research in sociology, while being methodologically advanced, fails to provide real explanations in which the mechanisms that produce observed outcomes are specified. For example, the relation between one’s educational level and one’s occupational status is often conceived as a simple positive linear relation: an individual from a certain social background receives a certain amount of education, and this education (among other individual characteristics) leads to a certain status position, and the sizes of these effects can be estimated empirically. The fact however that the individual with a certain level of education who is looking for a job is part of a social system in which other actors are also looking for good jobs is often ignored (Coleman 1984). One theory that at least pretends to provide an explanatory model of this stratification process and was therefore chosen as the main focus of this study is the theory by Boudon (1974). This theory basically consists of two models: the first concerned with the generation of inequality of educational opportunity (IEO), the second with inequality of status opportunity (ISO). The IEO model models the educational system as a series of transitions from one educational level to the next, where at each transitions students have to decide whether to proceed or leave the system, and their decisions are determined in direct ways (the “secondary effect”) and indirect ways via learning skills (the “primary effect”) by their social origin. The ISO model conceives of the labour market as a social system in which actors compete for the best status positions, and in which education is the major determinant of success. Together, these two models provide a simplified picture of the whole stratification process from social origin to a social status destination, via the educational system. Boudon used this theory, formulated as a numerical thought experiment, to investigate the relations between IEO, ISO and educational expansion, and found that educational expansion decreases IEO but also that a decrease in IEO does not necessarily result in a decrease of ISO.

In chapter 2, I discussed this theory in some more detail and looked into other theories and research related to it. Because Boudon’s original version of the theory was developed as a single numerical example, the validity of his conclusions based on this method was often doubted. Mathematical analysis of the IEO model however revealed that Boudon’s general conclusions are not very much dependent on the precise parameter values chosen by Boudon; in other words, his general conclusions do not only hold for his specific numerical example, but also more generally (Kosaka & Fararo 1976; Raub 1984). Simulation analysis of the ISO model on the other hand showed that Boudon’s general conclusions (the most important one being that ISO can remain stable while IEO declines) does only
hold for a limited range of parameter values (Müller-Benedict 1999). I also showed how Boudon’s (IEO-) theory relates in interesting ways to Mare’s (1981) theory on the relations between IEO, educational expansion and selection, that need further exploration. Like Boudon, Mare conceptualizes the educational system as a series of transitions in which the survival probability is determined directly and indirectly by social origin, but he also clarifies the way educational expansion influences IEO and selection processes in education. An interesting question is to what extent Mare’s analyses apply to Boudon’s models.

From the overview of empirical research related to Boudon’s theory I concluded that most studies only test Boudon’s general conclusions, without adapting the parameters of Boudon’s original model to the specific research situation. In other words, Boudon’s numerical example has often been taken too literal, with - in my opinion - the result that while the theory seems promising, it has not been tested adequately. Therefore I proposed to use a simulation approach to figure out the precise relations between inputs and outputs in Boudon’s models, and to use the results to assess how the theory performs in different situations.

In chapter 3, this simulation model was developed along with a set of research questions to give direction to the simulation process. With regard to IEO, the research questions concerned the relations between the strength of the primary and secondary effects, the number of educational transitions and educational expansion on one hand and IEO on the other hand. Another set of research questions focussed on the remaining association between social origin and learning skills at the highest level of education. According to Mare, this association would decline over transitions because of selection on skills, but the association at one transition would increase as a result of educational expansion. The questions is whether Boudon’s theory predicts the same. The research questions with regard to ISO aim at clarifying the relations between the inputs (IEO and the level of meritocracy), and the output ISO. Moreover, the effects of an extension of the original model – an increase of relative number of high status positions – is also examined. Then a simulation model was developed in which all the relevant parameters (the sizes of the primary and secondary effects, the number of transitions, educational expansion, the level of meritocracy and the proportion of high status positions) are systematically varied, using Boudon’s original numerical example as a starting point. The simulation resulted in a dataset with quantitative information on 84,000 different simulated “cohorts”; groups of virtual people with different stratification outcomes, depending on the model parameters. This dataset of cohorts is then analyzed (in chapter 4) by means of conventional statistical techniques in order to discover the relations between the model’s parameters and outcomes, and to derive hypotheses on these relations. The hypotheses are summarized in table 6.1.
Table 6.1: Hypotheses generated by the simulation model

<table>
<thead>
<tr>
<th>IEO</th>
<th>Result of empirical test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Given that there is an effect of social origin on survival probabilities, an increase in the number of branch points will result in an increase in IEO</td>
<td>?</td>
</tr>
<tr>
<td>2. Given that there are branch points, an increase of the direct effect of social origin on survival probabilities (secondary effect) results in an increase in IEO.</td>
<td>+/-</td>
</tr>
<tr>
<td>3. Given that there are branch points, an increase of the effect of social origin on school achievement (primary effect) results in an increase in IEO.</td>
<td>+/-</td>
</tr>
<tr>
<td>4. The larger the number of branch points, the larger the positive effect of the primary and secondary effects on IEO.</td>
<td>?</td>
</tr>
<tr>
<td>5. The larger the primary effect, the smaller the secondary effect, and vise versa.</td>
<td>+/-</td>
</tr>
<tr>
<td>6. The larger educational expansion, the smaller IEO</td>
<td>+/-</td>
</tr>
<tr>
<td>7. The larger the primary effect, the stronger the negative effect of educational expansion on IEO</td>
<td>+/-</td>
</tr>
<tr>
<td>8. The larger the secondary effect, the stronger the negative effect of educational expansion on IEO</td>
<td>+/-</td>
</tr>
<tr>
<td>9. The larger educational expansion, the stronger the effect of origin on achievement at the highest educational level</td>
<td>+</td>
</tr>
<tr>
<td>ISO</td>
<td></td>
</tr>
<tr>
<td>10. The higher IEO, the higher ISO</td>
<td>+</td>
</tr>
<tr>
<td>11. The higher the level of meritocracy, the higher ISO</td>
<td>+</td>
</tr>
<tr>
<td>12. The higher the level of meritocracy, the higher the positive effect of IEO on ISO.</td>
<td>+/-</td>
</tr>
<tr>
<td>13. When the level of meritocracy is high and IEO is low, an increase in the proportion of high status positions has a positive effect on ISO. When meritocracy is low and IEO is high, the effect of an increase of the proportion of high status jobs is negative.</td>
<td>+/-</td>
</tr>
<tr>
<td>14. The weaker IEO, the weaker the effect of education on status</td>
<td>+/-</td>
</tr>
<tr>
<td>15. The stronger meritocracy, the stronger the effect of education on status</td>
<td>+/-</td>
</tr>
<tr>
<td>16. The higher IEO, the weaker the effect of meritocracy on the effect of education on status</td>
<td>+/-</td>
</tr>
<tr>
<td>17. When the level of meritocracy is high and IEO is low, an increase in the proportion of high status positions has a positive effect on the effect of education on status. When meritocracy is low and IEO is high, the effect of an increase of the proportion of high status jobs is negative.</td>
<td>+/-</td>
</tr>
</tbody>
</table>

“+” = clear support, “-“ = clear falsification, “+/+” = mixed support, “?” = not tested

Some of these hypotheses are not very surprising, like hypotheses 1 to 4: it is almost trivial that stronger effects of social origin on the probability of school continuation leads to higher inequality, and that when these effects have more opportunities to do their work (more branch points) IEO also becomes higher. The simulation results however do not only show that such is the case, but do also give an idea of the quantitative sizes of these effects. Hypothesis 5, on the interaction between primary and secondary effects, is a little surprising. Although one of the aims of simulation research is to produce counterintuitive results, it is hard to imagine why primary and secondary effects would have an effect on each other. I suspect this is an “artefact” of the simulation in the sense that it is the result of the way primary and secondary effects are implemented in the simulation (see equations 3.2 and 3.3), but which does not follow from the assumptions of the model.

The hypotheses on expansion and IEO (6-9) show that Boudon’s model generally reproduces Mare’s hypotheses on expansion and selection. Theoretically, IEO declines as a result of expansion, independently of the primary and secondary effects (H6), and expansion also leads to an increase of the association between social origin and learning skills after selection, or “conditional primary effect”
Hypothesis 7 and 8 however postulate relations between the primary and secondary effects and expansion that were not foreseen.

Since one of the main goals of this study (as stated in chapter 2) was to provide a more thorough empirical test of Boudon’s model using the simulation results, I made an attempt to do so in chapter 5. In the case of the IEO model however, I must conclude that this attempt was not fully successful. The main reason for this was that the testing of relations on the cohort level as postulated by the hypothesis requires information not often available in large scale empirical data, most importantly information on learning skills. In this case, only four school cohorts with this information were available, which made “conventional” testing by estimating the expected relations impossible. An alternative way was found by comparing the predicted trends of the dependent variables to the trends that would be predicted by the simulation model, based on the values of the independent variables (primary and secondary effects, expansion etc) found in the data. By this method, historical trends in IEO-related outcomes in the Netherlands were compared to the trends that would be predicted by the model, given the observed parameters. The disadvantage of this method is that the hypotheses can only be tested as a system, and not separately: if there is a difference between the observation and the prediction, it is very hard to say which one of the predicted effects goes wrong. On the other hand, it must be stressed that with a coherent model such as the simulation model used here, hypotheses should strictly speaking only be considered as a system, because they are all derivations of the same (relatively small) set of assumptions underlying the model.

Unfortunately, there did not appear to be much variation in the independent variables, which made testing the hypothesis even harder. Since the number of educational levels was constant, hypotheses related to this variable could not be tested at all (H1 and H4; see table 6.1). With regard to the other hypotheses, support was mixed. On one hand, the model seemed to be able to predict the observed trends in IEO and the association between social origin and skills at the highest level fairly well. In the latter case, the model was also able to explain the upward trend in this dependent variable by the developments of two different independent variables: at first the increase is caused by educational expansion, and later by an increase of the primary effect itself. On the other hand, the model was also wrong in a number of cases (for example while predicting a very slight increase in IEO while in reality there was a slight decrease). Given the fact that there is very little variation in the independent variables, a strict test of the hypotheses would consider them falsified. But given the small number of observations used here I propose to grant the model the benefit of the doubt.

The hypotheses on ISO are probably somewhat more interesting than those on IEO, because the model they result from is more complex, and the hypotheses are more counterintuitive. Hypotheses 10 and 11 confirm the earlier (theoretical) findings by Müller-Benedict (1999), and refute Boudon’s assertion that ISO is not affected by a decline in IEO (Müller-Benedict already showed that such is only the case
for a limited range of IEO values). Hypothesis 13 is in my opinion the most surprising of the results of the simulation. It is often suggested that Boudon’s finding that ISO can remain stable while IEO declines is caused by a discrepancy between the development of the educational distribution (more highly educated people) on one hand and the job structure on the other hand: the effect of more educational equality is neutralized because the job structure does not adapt to the distribution of education, or is at least moving at a slower pace. The perfectly rigid job structure in Boudon’s original model was also criticized for being highly unrealistic (Hauser 1976). The simulation study however, in an extension of Boudon’s original model, showed that even when this “unrealistic” assumption of a rigid job structure is relaxed, this does not necessarily lead to more status equality. On the contrary, under certain circumstances (when IEO is low and meritocracy is high), an increase of the proportion of high status positions can even lead to an increase of ISO. This means that theoretically, the stability of ISO under a decline of IEO could be the result of a change of the job structure in direction of more high status positions rather than being the result of a discrepancy between the educational distribution and the job structure, because the change in the job structure could offset the effects of declining IEO. The hypotheses on the returns of education (14-17) are similar to the ones on the development of ISO, and suggest that in the model the association between social origin and status (or ISO) is mediated by the effect of education on occupational status.

The hypotheses on ISO and the returns of education were tested empirically using data from the International Social Mobility and Politics File (Nieuwbeerta & Ganzeboom 1996). This dataset contains the combined information from a large number of surveys collected in 16 different countries between 1956 and 1991, and thus provided an opportunity to test the hypotheses on ISO using much more observations than where available for IEO. At first, the same procedure as for IEO was applied to the ISO hypotheses: the trends predicted by the model where compared to the observed trend, while focussing on labour market cohorts in the Netherlands between 1946 and 1980. To operationalize the level of meritocracy, the unemployment rate was used; the other independent variables (or model parameters) could be more or less directly estimated from the data. It appeared that the trends in ISO and the returns of education can be predicted with the help of the simulation model quite well, although the model consequently underestimates the level of ISO and overestimates the effect of education. A plausible explanation for this difference is that the simple model that was implemented in the simulation model assumes no direct effect of social origin on status. If in reality there is such an effect (which is very likely), then ISO might be higher because there an extra direct effect of social origin is added to the indirect effect through education. Interestingly however, the difference between the observed and predicted values is almost constant, such that the regression model used to generate the predictions could be easily modified by adding a constant to match the observations. This conclusion is however complicated by the results with regard to the returns of education (or the effect of education on status): here the model predicts a general decline, while in reality the effect is more or less stable over time.
Another observation was that in many cases, the predicted trend seemed to be somewhat ahead of the observed trend, which suggests that changes in the independent variables need more time to have an impact than the model assumes.

Although this method of testing the model by comparing predicted and observed trends is instructive, it does not allow us to test hypotheses separately. A method which does allow for testing the hypotheses separately is estimating expected effects from the data, which is the way hypotheses are most commonly tested in the social sciences. In the case of IEO, not enough observations were available to do this, but for the ISO hypotheses, enough observations could be collected by using four more countries (in addition to the Netherlands) from the International Social Mobility and Politics File, such that at least main effects could be estimated. In this way, hypotheses 10 & 11 could be directly confirmed: there are positive effects of IEO and the level of meritocracy on ISO. Moreover, the analysis also showed that the proportion of high status positions can have a positive effect on ISO, which is at least partial support for hypothesis 13. For further testing of the interaction hypotheses, the number of observations was still not sufficient. The fact however that the model from which these hypotheses are derived is able to predict trends in ISO with some accuracy can again be considered as indirect support.

What do these results tell us about the value of Boudon’s theory, and more importantly, about the mechanisms behind stratification processes? I have already indicated that my aim to provide a proper empirical assessment of the value of Boudon’s model has only been partly achieved: most hypotheses could not be conclusively confirmed or falsified. Especially in the case of the IEO model, one could well argue that the evidence is more against than in support of the hypotheses. However, the simulation analyses showed that the way the process of educational stratification is modelled here – assuming primary and secondary effects in a sequence of transitions, and expansion through a ceiling effect - is in its outcomes more or less consistent with an influential theory like Mare’s. Moreover, the simulation model provides more accurate (quantitative) information on the expected effects, which in the case of the conditional primary effect even proved to be fruitful in predicting and explaining observed trends.

For the ISO model support was generally stronger: observed trends in ISO and the returns of education could be predicted fairly well, while some hypotheses could also be confirmed more directly. The results suggest that ISO does indeed at least partly depend on IEO, which is in line with the hypotheses but contrary to Boudon’s own assertions. It is important to note that the trends in ISO were predicted using a simple model in which direct effects of social origin or “dominance” effects play no role. ISO is indeed underestimated, but at a constant rate, which suggest that if there are any dominance effects, they have been constant in the observed period in the Netherlands and did not have much effect on differences in ISO between years. If it is accepted that the simulation model used here
makes predictions that are sufficiently accurate, then this implies that we might not really need dominance effects to model ISO. It that case, it would be sufficient to assume that the process that leads from a certain level of education to a certain status position takes place through a mechanism of competition in which the level of education is an important “weapon”, and where the strength of that weapon (meritocracy) is variable (influenced by, for example, the economic situation). In the end, the outcome of the process depends on the distribution of competitors (IEO) and the value that is attached to education (meritocracy), but also on the distribution of available positions; although the hypotheses on this effect were not fully tested, the analyses at least showed that the role of the status distribution is far from trivial.

Overall, I think that the results show that although Boudon’s models are not in all cases supported, they do provide a theoretical framework for stratification processes that can be fruitfully applied to explain trends in educational and occupational inequality. Using very simple assumptions that model the behaviour of the relevant actors and the social structure they operate in, the models produce – with the help of a simulation approach as applied in this study – testable predictions that can be reasonably compared to outcomes observed in reality. This is already a significant accomplishment for a theory that was (in an influential review) dismissed for being clearly unrealistic and “sociologically naive” (Hauser 1976: p. 919; 912). The fact that in some cases (especially related to ISO) the theory appears to be able to approximate observed trends fairly well suggests that these models (contrary to Hausers’s contention; p. 923) indeed do deserve wider application.

6.2 The merits of simulation

A secondary goal of this study was to explore the possibilities of the use of a simulation approach in stratification research. I think that the results as summarized above have shown that simulation has some important advantages.

In the first place, simulation can reveal relations between variables that are sometimes quite counterintuitive and were not expected beforehand, like the effect of the proportion of high status positions on ISO. This is especially helpful for models that are relatively complex, like the ISO model. In the second place, the simulation has allowed us to clarify the consequences of a model like Boudon’s such that they could be applied to different situations and thus be tested empirically. In previous research, Boudon’s conclusions were too often taken as generally applicable (which is perhaps also the way they were meant by Boudon) and applied to specific situations they did not apply to. The use of simulation makes it easier to see what the implications of a theory would be in different situations, and to test it accordingly. Although the empirical test in themselves were not always successful, the simulation model was at least able to “reformulate” the theory in such a way that it can be properly tested, (given that the right data are available,) because it helped to generate more precise
hypotheses on the predictions of the theory in different empirical settings. I think this is a step forward compared to previous research related to Boudon’s models.

A third, related point is that simulation makes it possible to derive hypotheses that are formulated quantitatively. In most sociological theory, hypotheses are formulated in general terms, and postulate relations like “the larger $x$, the smaller $y$”. In contrast, the simulation model used here generated “hypotheses” formulated as regression equations, that were accurate enough to predict trends in inequality year by year, and compare them in a reasonable way to observed trends.

There are however also problems with simulation. The first problem is that, exactly because simulation provides theoretical results that were not easily seen beforehand, there is always a risk that results are not really following from the assumptions of the model, but instead from the way they were (perhaps wrongly) implemented in the model, and are in a sense artefacts. In this study for example, this is possibly the case with the finding that there is an interaction effect between primary and secondary effects. Of course, this problem can be avoided by good programming, but exactly the advantage of simulation that it can make sense of something the researcher did not fully understand beforehand also makes it vulnerable to flaws.

A second problem is that the quantitative relations resulting from a simulation are not always easily interpreted, when it is not clear what the exact meaning is of the values used. For example, the simulation model used here uses a parameter for meritocracy, but the exact meaning of a “meritocratic parameter” with a value of 70% is unknown. This can make it hard to test hypotheses generated by a simulation model, especially when absolute values matter. In this case, the problem could be more or less satisfactorily solved by trying different scales for meritocracy because meritocracy was the only variable with this problem. When the substantial meanings of the values of more than one parameter are unclear, the problem might be much more serious.

Lastly, there is the (somewhat philosophical) problem of the validity of predictions from a (over)simplified model. Because simulation models are usually quantitative models, they must use simplifying assumptions that are sometimes obviously very unrealistic. The assumption in the theory according to which people who start their educational career at the same moment also enter the labour market at the same time is a good example: in reality this is obviously not the case, because obtaining a higher educational degree usually takes more time, during which others who left school at a lower level already enter the labour market. The distribution mechanism that was used here in which there is no direct influence of origin on status attainment is another example. Yet, the predictions from the model that are compared to reality do result from this unrealistic representation of reality. On the one hand, models are supposed to be simplifications of reality (which is why they are models); on the other hand, it seems problematic to agree with predictions that follow from a theory that is obviously wrong, even if the predictions appear to be correct. This problem is too complex to solve here, but should at least be mentioned.
6.3 Suggestion for further research

I conclude this study with a few remarks on possible research efforts that could improve or extend the results of the study described here.

A first set of improvements concerns the empirical testing of the hypotheses resulting from the simulation model used here. Above, I explained that some hypotheses could not be tested satisfactory or not at all, because of the availability of data. An obvious improvement would be to find the right data and test the hypotheses more thoroughly. This could involve finding data in which the number of educational levels is varied, using more educational cohorts, or applying the analysis as applied to the Netherlands to other countries as well.

Second, the simulation data used here allow for more analyses than I have done here. The relation between different measures of IEO and ISO (odds ratios or disparity ratios versus regression coefficients) for example deserves more attention. Also, the analysis could be carried out in more detail: since the simulation model in principle generates data on the individual level, many types of analyses are possible. For example, it would be possible to apply log-linear regression techniques in the way indicated by Mare (1981).

A last set of suggestions concerns improvement and extension of the simulation model and the underlying theory itself. One could study more different parameters (like the number of status levels), introduce dominance effects or try to find a new method of distribution that is more easily accepted as realistic, and in which ascription is also systematically modelled. Another interesting question (already mentioned by Fararo and Kosaka (1976)) is the relation between the ISO outcomes and the decisions in education made by younger generations. Since these decisions are assumed to result from cost and benefit analyses with regard to obtaining a high status position, it can be expected that the relation between education and status that results from the competitive meritocratic process and is observed by the actors has some effect on their decisions. To model this relation would however be quite a challenge. Further integration of formal decision models (like the one proposed by Breen and Goldthorpe (1997)) and mobility models like the job competition model used here would be helpful in this regard.
Literature


