Academic specialization choices and academic achievement
prediction and incomplete data
Smits, N.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
8 Overview and conclusions

This dissertation has two main themes: university students' choice of a specialization, and university students' academic achievement. A prerequisite for performing prediction studies on these two topics is dealing with missing data. Therefore, this dissertation addresses missing data analysis as well. In economics, specialization choices are explained by the financial benefits of education. In psychology, specialization choices are mainly explained by the vocational interest in the study content of academic specializations. This restricted set of benefits of education was expanded with a number of other benefits, such as the effort a specialization requires, and the possibility to work with other people. The extended set of benefits proved to be fruitful for predicting academic choices. In different branches of the behavioral, and social sciences, academic achievement is explained using different sorts of variables. In chapter 2, academic achievement was explained on the basis of pre-university (VWO) grades. In chapter 4, four blocks of variables – a background, an economic, a time budget, and a psychological block – were compared at their usefulness of predicting academic achievement. Large differences in the usefulness between the blocks were apparent. The psychological variable block gave the best and most stable predictions. In psychological and educational research, incomplete data are very often encountered. In this dissertation two missing data problems were addressed. In chapter 2, incomplete VWO grade files were considered a missing data problem. Several missing grade methods were compared at their theoretical and empirical differences, and at their performance in two prediction of academic achievement analyses. The second type of missing data were the missing item scores resulting from the application of incomplete test designs on educational and psychological tests. In chapter 3, several methods for the estimation of incomplete item scores were compared on the reconstruction of test reliability, and test predictive validity. The different methods produced good results at either reliability or predictive validity estimation, but not at both. It was concluded that measurement precision and predictive validity of incomplete tests could not be optimized simultaneously. One method that proved to be successful at reconstructing predictive validity was applied to the scales in the psychological variables block in the survey of chapter 4.

8.1 Incomplete data

In psychological and educational research, missing data are often encountered. In this dissertation, two sorts of missing data are distinguished: missing data that
are beyond, and missing data that are not beyond the control of the researcher. An example of the first sort of missing data is an unintendedly skipped item in a psychological questionnaire. An example of the second kind of missing values are data that, on the basis of efficiency considerations, are intended not to be collected (see, e.g., Schafer & Graham, 2002). In this dissertation, several methods for dealing with missing data were compared.

In chapter 2, the Grade Point Average (GPA) ('gemiddelde eindexamencijfer' in Dutch) was considered a technique for missing grades. In study 1, theoretical and empirical differences between GPA, and seven alternative methods for unavailable grades were studied. All methods replaced, or 'imputed', unavailable grades with an estimate. The seven alternative missing grade techniques were subject mean substitution, corrected subject mean, subject correlation substitution, regression imputation, EM algorithm imputation, and two multiple imputation methods—Stochastic Regression Imputation (SRI), and Data Augmentation (DA) procedure NORM (Schafer, 1997, 1998). The missing grade techniques were very different. DA and SRI appeared to be superior as missing grades techniques. In study 2, completed grade records were used in two analysis in which academic achievement was predicted. One analysis was based on unweighted grades, and the other was based on weighted grades. In both analyses, alternative methods produced better, and more stable predictions of academic achievement than GPA. It was concluded that some alternative missing grade methods were better than GPA.

In chapter 3, the application of incomplete test designs to psychological and educational questionnaires was studied. As a result of the recent availability of advanced techniques for dealing with missing values, data can be collected using an incomplete design, and data that are not recorded can be estimated. Applying incomplete designs to questionnaires can be used as a way to increase response rates in survey research. As shorter questionnaires decrease the respondents' burden, response rates can increase. The usefulness of incomplete designs has especially been studied for survey questionnaires. However, psychological and educational tests are different from normal surveys, because they need more items to measure constructs. Developers of educational and psychological tests are commonly faced with two conflicting test goals (see, e.g., Lord & Novick, 1968). First, a test should be valid for the prediction of a given criterion, such as academic achievement. Second, the test should be a precise measurement of a given characteristic of an individual, such as achievement motivation. When incomplete test designs are applied to psychological and educational tests, the test goal should be taken into account, because the available methods for missing item scores seem to meet only one of these test goals at the expense of the other. Two sorts of model based techniques for dealing with incomplete test data exist. The first group of techniques are measurement models like Item Response Theory (IRT), which allows for the estimation of a common trait for all examinees on the basis of answers to different sets of items. The second group of techniques are models, such as DA, that estimate missing items scores on the basis of all the data that were recorded, a criterion included. In this study, paper and pencil tests were simulated, and a third part of the item scores was deleted using a blocked interlaced
design. Next, two IRT methods, two versions of DA and two simple techniques for missing item scores were applied to impute the deleted item scores. The methods were compared at their ability to reconstruct test scores, reliability coefficients, and predictive validity coefficients. All methods gave good reconstructions of test scores. IRT gave best results of the reconstruction of reliability, and DA gave the best estimation of predictive validity. None of the methods could simultaneously reconstruct accurately measurement precision and predictive validity.

8.2 University students' specialization choice

In some theories of academic choices it is assumed that students choose a specialization in order to meet a certain goal. In the economics of education, academic choices are explained by financial returns to education (see, e.g., Vella & Gregory, 1996). Students choose the specialization from which they gain most. In psychology, specialization choices are explained by the interest in the study content of specializations (see, e.g., Holland, 1997). Students choose the specialization with the study content that best fits their interests. In this dissertation it is argued that this pair of benefits of education gives a too limited view of what university education can yield. The set of benefits of education was expanded with other sorts of benefits, such as the effort a specialization requires, the possibility to work with other people, and the opinion of parents.

In chapter 5, freshmen of nine different university courses (law, history, economics, psychology, political science, English linguistics, Spanish linguistics, medical biology, and dentistry) participated. First, the students reported the expected income after graduation of their own specialization, and of other specializations. Second, they evaluated thirteen benefits associated with academic disciplines. Students from the nine disciplines differed considerably in the extent to which they anticipated maximum financial earnings. In addition, large differences were found in the importance ratings of the 13 benefits. The evaluations and gender were used in a multinomial logit model in order to predict academic specialization choice. The model had a rate of correct classification of 37%. Further analysis showed that omitting gender from the model did not result in a substantive downswing in the predictions.

In chapter 6, the set of 13 benefits was applied again. In this study, students from law, psychology, economics, and psychology participated. Like in chapter 5, evaluations of the importance of benefits were used. In this chapter, these evaluations were called 'desires'. In addition, students were assessed at their perceptions on the 13 benefits of the specialization they had chosen, and the three remaining specializations. Students from the four academic disciplines differed both in their desires and their perceptions. Five methods were applied to predict academic choices: discriminant analysis, the Conditional Logit Model (CLM), Multi-Attribute Utility Theory (MAUT), Equally Weighted Criteria (EWC), and the Euclidean distance. These five methods made a different use of students' desires and perceptions: discriminant analysis exclusively used the desires; CLM
Overview and conclusions

and ECW made only use of the perceptions; MAUT and the Euclidean distance made use of both the desires and the perceptions. All methods gave reasonable predictions; CLM was the best method, the discriminant analysis was the least adequate method. The set of benefits turned out to be very useful for studying specialization choices. In addition, the design of assessing both general desires and perceptions of specific specializations appeared to be very fruitful.

In chapter 7, two of the methods that were used in chapter 6, were compared. The first method, CLM, comes from economics, and is a model for discrete choices. The second method, MAUT, originates from psychological decision theory, and is a prescriptive method for individual decision makers. Both models make use of evaluations of attributes that are specific for specialization choices. However, CLM estimates population attribute weights, whereas MAUT uses attribute weights that are determined by the decision makers themselves. Both methods were compared in an analysis in which the specialization choices of two cohorts of university freshmen were used. The two models gave good classifications of the students, and had a high overlap in classification. Moreover, the attribute weights of the two methods appeared to be related. It was concluded that CLM and MAUT were quite similar, both in theory, and in their empirical results.

In all the chapters associated with the specialization choice of university students, missing values encountered in the assessments of the desires and perceptions were replaced using missing data procedure NORM.

8.3 University students’ academic achievement

The academic achievement of students in higher education has been studied in many branches of the behavioral and social sciences. Many different variables have been used to predict academic achievement. The predictor that has been used most often is intelligence. However, in the Dutch situation, the predictive power of intelligence has always been low (see, e.g., Busato, 1998). Other researchers have stressed the achievement in previous education, commonly measured by school grades. Some have emphasized personality (see, e.g., Busato et al., 2000). It has often been reported that students who are conscientious have higher achievement than students who are not (see, e.g., Wolfe & Johnson, 1995). However, psychological variables are not the only variables that have been used to explain academic achievement. For example, basic personal characteristics, such as gender and age, have been applied. Female students have been reported to perform better than their male colleagues in many courses (see, e.g., Mau & Lynn, 2001). Sometimes, the time that students allocate to several activities, such as sleeping, working, and studying is used to explain academic achievement. The time that is allocated to studying is not always positively related to academic achievement (Olivares, 2002). Some students have explained academic achievement from the financial situation of the student. Students who expect more financial benefits from their specialization, have been reported to be more persistent in their study (see, e.g., Oosterbeek & Webbink, 1995).
In chapter 2, pre-university (VWO) examination grades were used to predict university freshmen's academic progress after one year. The correlation between VWO grades and study progress was about .30.

In chapter 4, the academic achievement of university freshmen was predicted on the basis of four blocks or variables: background variables, economic variables, time budgets, and psychological variables. The psychological variables block consisted of four sub-blocks: personality, motivation, self-efficacy, and study skills. The knowledge on incomplete designs that was acquired in chapter 3, was used in this chapter. To the scales of the psychological variable blocks an incomplete design was applied. As a consequence a third part of the item scores was missing for each student. This missing part was estimated using NORM.

The usefulness of the predictive blocks was assessed on the basis of multiple $R$ coefficients, and cross validated multiple $R$ coefficients of the regression models that applied the blocks to predict academic achievement. Two measures of academic achievement were used: GPA, and study progress. Chapter 2 showed that GPA was not the best method for dealing with unavailable grades. However, in this study, methods that appeared to be superior in chapter 2 gave computational problems, and GPA had to be used. In this study, students from economics, psychology, communication, and medical science participated. The academic disciplines differed to a large extent in what blocks produced good and consistent prediction results. Generally speaking, the psychological block gave best results, for both measures of academic achievement. So, in spite of the presence of incomplete item scores in the survey, completed psychological tests gave good results. In addition, the academic achievement of psychology students was more easily predicted than the achievement of students in other specializations.