Issues in growth curve modeling
Stoel, R.D.

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: http://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.
Introduction

A broad range of statistical methods exists for analyzing data from longitudinal designs. Each of these methods has specific features and the use of a particular method in a specific situation depends on such things as the type of research, the research question, and so on. The central concern of longitudinal research, however, revolves around the description of patterns of stability and change, and the explanation of how and why change does or does not take place (Kessler & Greenberg, 1981, p.1).

The spectrum of longitudinal methods varies among the different disciplines of social science. It ranges from the panel, or wave, design in sociology, to age-related repeated-measurement methodology in developmental psychology, to various forms of single subject designs in operant psychology, and to time-series arrangements in econometrics and political sciences. However, the central focus in all of these instances is the ordered study of processes. Although the definition of process varies by discipline and emphasis, the common denominator is that process research centers on how and why a phenomenon behaves in time in relationship to both constancy and change (Baltes & Nesselroade, 1979, p.2).

One of the most common designs for longitudinal research in the social sciences is the panel design, in which a sample of cases is observed at more than one point in time. While in a time-series design one subject (or a few) is measured on many occasions, and in a cross-sectional design more subjects are measured on one occasion; panel designs are essentially a combination of both time-series and cross-sectional designs (Rogosa, 1979). An important feature of a panel design is that, in contrast to static cross-sectional designs, change is explicitly incorporated into the design so that individual (or other unit-level) changes in a set of variables are measured directly (Finkel, 1995, p.1).
Panel designs have been in use already for many years in social science. Psychometricians, for example, used panel designs to study the reliability of their measurements; developmental psychologists used panel designs for studying the development of children at various ages; marketeers used panel designs to model the changes of brand preference in marketing studies; and panel studies were applied in the study of change in political preference. Many of these studies were inspired by the work of Paul Lazarsfeld (e.g. Lazarsfeld & Fiske, 1938; Lazarsfeld, 1948). From his work, Lazarsfeld concluded that, in comparison to cross-sectional and trend designs, the panel design had to be preferred in all aspects.

Although the concept of panel designs and the understanding of their possibilities actually dates back to the beginnings of modern-day survey research, it took until the late 60's for the design to gain more interest among social scientists. This growing interest in panel designs can be explained partly by the growing interest in developmental change and partly by the more common use of structural equation methods for analyzing 'causal' relations between variables. By that time social scientists came to understand that structural equation models could only reach conclusions about dynamic processes on the basis of cross-sectional data under very restrictive circumstances; and since panel studies were thought of as being able to investigate causal relationships, free from a priori assumptions, the analysis of panel data with structural equation methods seemed highly promising in establishing causal relationships. Though it was shown quickly that this was not really the case, methodologists did find linear panel models extremely useful in specifying the processes involved in social scientific research.

Several distinct statistical techniques are available for the analyses of panel data. One of these techniques is growth curve analysis. Growth curve analysis has been said to overcome some of the limitations of more traditional approaches to the assessment of change, and it is currently receiving a lot of attention in the social and behavioral sciences. Growth curve analysis is a statistical technique to estimate the parameters, or latent factors, which represent the growth curves that are assumed to have given rise to the structure of the repeatedly measured outcome variable over time. Growth curve analysis can be applied to get a (unconditional) description of the mean growth over a certain period of time. However, the emphasis of the technique lies in explaining the differences between subjects of the parameters describing their growth curves; in other words, in the systematic inter-individual differences in intra-individual change.

This dissertation focuses on issues in the analysis of continuous outcome panel data using growth curve analysis. The aim of the present work is to enhance the method of growth curve analysis, and to encourage its adequate application, by comparing two generally accepted analytical techniques for growth curve modeling (Chapter 1), by a detailed investigation of selected extensions of the basic growth curve model (Chapter 2, 4 and 5), and by detailed application of growth curve analysis to empirical data (Chapter 3), followed by a discussion of the results in Chapter 6. It is, with great emphasis, not the purpose of this dissertation to contrast growth curve modeling to alternative statistical approaches for the analysis of panel data, in spite of the fact that such a comparison might prove a fruitful (future) activity. It is, therefore, assumed in the following that the growth curve
model is an adequate candidate to represent the structure of the continuous outcome variable over time.

A major reason to limit the discussion in this thesis to growth curve analysis is a practical one. Growth curve analysis has been, and will be, frequently applied in the social and behavioral sciences for the analysis of panel data. Thus, without questioning the appropriateness of growth curve analysis and the growth curve model in general, there is a great number of social scientists to which growth curve analysis provides a useful way to investigate their research questions. It is stressed that the present focus does not imply that growth curve analysis is the current best practice for all longitudinal investigations, nor that it is implied in any situation at all. The fact that a (set of) technique(s) receives so much attention in the literature calls for a thorough study of the techniques themselves.

In Chapter 1 the focus is on two techniques for the analysis of panel data; longitudinal multilevel regression analysis (Bryk & Raudenbush, 1987, 1992; Goldstein, 1986, 1987, 1995) and latent growth curve analysis (McArdle, 1986, 1988; Meredith & Tisak, 1990; Willett & Sayer, 1994). The former takes a multilevel regression perspective, and the latter a structural equation modeling perspective. Often, the choice between longitudinal multilevel regression analysis and latent growth curve analysis is not guided by pure substantial and/or methodological arguments; the same holds for the choice between growth curve analysis and alternative data analytic approaches. Cross-sectional multilevel regression analysis, for example, is quite popular in the educational sciences. Consequently, a researcher coming from this field will be attracted to longitudinal multilevel regression analysis when confronted with a longitudinal data set because multilevel analysis is the current practice in his or her field, and this is the technique that he or she is familiar with. Even if longitudinal multilevel regression analysis could be the current best practice in a specific situation indeed, this can only be concluded if alternative techniques are considered as well. Likewise, it is not difficult to think of a similar example with respect to the decision of whether growth curve analysis is the appropriate technique. The techniques in this chapter will be compared in their flexibility to produce possible extensions of the basic growth curve model. This chapter is also the most general chapter of this dissertation, and it creates the context in which the further discussion will take place. For this reason it is placed at the beginning of the dissertation. The other chapters are presented in the order in which they were written.

Chapter 2 investigates consequences of linear transformations of a key factor of the latent growth curve model\(^1\): the time scale. Although it may not be obvious at first sight why timescales should be transformed, such transformations occur implicitly if the theoretical start of a process and the first measurement occasion do not match. The idea of writing this chapter comes from the observation that the results from a multilevel regression analysis using a MLwiN software package (MLwiN1.10---Rasbash, Healy, Browne & Cameron, 1998), did not converge to the results from several structural equation modeling programs. The reason for the

\(^1\) Chapters 2, as well as Chapter 4, take the latent growth curve model as point of reference, while the consequences also hold for the longitudinal multilevel model.
nonconvergence of these seemingly similar models is that MLwiN uses, by default, one time unit prior to the first measurement occasion as the origin of the process, and this approach is not common practice in structural equation modeling. Although consequences of timescale transformations for growth parameters gain ground by methodologists and applied researchers the question remains what the consequences would look like for effects of exogenous covariates on the growth parameters. Chapter 2 investigates this question.

Chapter 3 investigates the relations between trajectories of school related concepts by means of a multivariate growth curve analysis. In the final model, the developmental processes of self-confidence, school-investment, and language acquisition of children during their elementary school period are related to each other, as well as intelligence. This model provides an extension of the basic latent growth curve model in the sense that it incorporates the factor structure of the repeatedly measured variables, the possibility of nonlinear growth, and it incorporates a time invariant covariate. Reconsidering the analysis procedure we became aware of possible problems with the use of these extensions of the basic latent growth curve model.

In Chapter 4 it is shown that the standard way of incorporating time-invariant covariates in a latent growth curve analysis may be unnecessarily restricted in certain instances, and a more general way of modeling the time-invariant covariate will proposed. Chapter 5 investigates opportunities and pitfalls of the nonlinear latent growth curve model, and the higher order growth curve model. It concentrates on the topic of longitudinal measurement invariance in reaching conclusions about ongoing processes. The final Chapter 6 presents an integration of the previous chapters.

As is becoming current practice in the social and behavioral sciences in The Netherlands, all chapters of this thesis have been published, accepted or submitted for publication in international journals. Although there is order in the way they are presented in this thesis, the chapters are self-contained and can be read separately. The chapters are kept as similar as possible to their published or submitted counterparts, causing a small variation in notation between the chapters, and some overlap in content. I hope that this will not affect the readability of the thesis in its entirety.