

RECENT PHD THESES

University of Hasselt (UHasselt)

Jacobs Tom. *Non-linear mixed-effects modeling for complex biopharmaceutical data* (2009) – Promotor: Pr. G. Molenberghs.

A lot of the attention on statistical models used to be focused on (generalized) linear models. However, several phenomena in nature exhibit nonlinearity. One well-known example is the plasma concentration-time profile after the administration of a drug product. With the increase in computer capacity and the availability of software such as nonmem, SAS, WinBugs, and monolix, fitting the nonlinear mixed effects models required in pharmacokinetics became feasible. As the plasma concentration-time profile is the driving force behind the efficacy and safety of a drug product, a good understanding and an accurate estimation of the models is required.

Controlled-release formulations pose additional difficulties: not only the metabolization and the elimination determine the model fit, but also the accuracy and precision of the in-vivo release mechanism ought to be investigated. However, the in-vivo release is not observed. Therefore, one combines a multiple of responses and models to overcome this. Gillespie and Veng-Pederson (1985) demonstrated that the controlled-release plasma concentration-time profile can be modeled as the convolution product of the immediate release plasma concentration-time profile and the in-vivo release of the drug product. As one is not interested in the elimination of the controlled-release formulation, this is estimated from the immediate-release plasma concentration-time profile. A hypothetical relation between the in-vitro and in-vivo release is imposed (O'Hara et al. 2001) and the adequacy of the model fit for the controlled-release plasma concentration-time profile is assessed. Such a model is referred to as an in-vitro – in-vivo correlation model (IVIVC). These models play a crucial role in the pharmaceutical industry: once an IVIVC is established, it allows assessing the impact of changing to the manufacturing procedures, batch differences, etc. based on an in-vitro test. IVIVC models also enhance formulation development.

In the dissertation, a model was presented that copes with heterogeneous formulation, i.e., with a dual release mechanism. The methodology was also extended to a one-stage procedure to allow the exchange of information

between the different sub-models of the IVIVC, coping with the in-vitro release time profile, and the immediate-release and controlled-release plasma concentration-time profile. Further, the existing model diagnostics (%PE) imposed by the authorities (FDA 2003) were criticized and local influence was introduced to detect potential outlying plasma concentration-time profiles. Further, the range of applications was extended: the IVIVC model was combined with a PK/PD model to link the in-vitro dissolution properties of a drug formulation to the in-vivo receptor binding. As such, changes in the in-vitro dissolution profile were translated into changes in the clinical effects of the controlled-release formulation. This allows determining clinically significant dissolution specifications.

In the same vein, the existing bioequivalence methodology, which determines whether two tablets or capsules yield the same drug exposure, was modified to incorporate the therapeutic window of the drug product; the newly proposed bioequivalence acceptance ranges are more conservative when minor changes in exposure lead to major clinical effects, whereas more liberal acceptance ranges are imposed when minor clinical changes are observed for large changes in exposure.

U.S. Food and Drug Administration, Center for Drug Evaluation and Research; "Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drug Products --- General Considerations"; 2003.

Gillespie, W.R., Veng-Pedersen, P. (1985). Gastro-intestinal bioavailability: determination of in vivo release profiles of solid dosage forms by deconvolution. *Biopharmaceutics and Drug Disposition*, **6**, 351–355.

O'Hara, T., Hayes, S., Davis, J., Devane, J., Smart, T., and Dunne, A. (2001). in vivo—in vitro correlation (IVIVC) modelling incorporating a convolution step. *Journal of Pharmacokinetics and Pharmacodynamics*, **28**, 277--298.

Laenen Annouschka. *Psychometric Validation of Continuous Rating Scales from Complex Data* (2008) – Promotor: Pr. G. Molenberghs.

Psychometric Validation of Continuous Rating Scales from Complex Data
In clinical trials, the response of interest is sometimes difficult to measure. This happens in psychopharmacological trials, where interest lies in measuring traits like anxiety or depression, but also in other fields, when measuring concepts like quality of life or pain. Rating scales are generally used to measure this type of latent variables. Such scales then need to be *valid* and *reliable*. A scale is valid if it actually measures the latent variable that we are aiming at, and it is reliable if the measurement error is limited. However, the latter is not a fixed scale characteristic but it is population dependent. The same scale applied to two different populations can result in different reliabilities. In our work we have focussed on the question how we can evaluate the reliability of the outcome measurements within a clinical study.

The classical psychometric approach for evaluating reliability is based on a very restrictive modelling framework that is unlikely to hold for longitudinal clinical trial data. We therefore base our approach on the flexible family of linear mixed models that can handle the typical characteristics of repeated measurements. The first step was then to extend the concept of reliability as defined in the classical test theory to this more general setting. For doing this we used an axiomatic approach and proposed four properties that any measure for reliability should fulfill. These properties state that any measure for reliability (1) should always lie between zero and one, (2) should be zero only if there is only measurement error, and (3) should be one only if there is no measurement error at all, and finally (4) should be equal to the classical definition whenever the conditions of the classical theory are satisfied.

The above definition led to a whole family of reliability measures of which all members are weighted sums of the same basic elements: the roots of the equation $q(\lambda) = |\Sigma - \lambda V| = 0$; where Σ expresses the error variability and V the total variability. Two measures, R_T and R_Λ , were of particular interest and have been further scrutinized. In a longitudinal context, the R_T coefficient expresses the average reliability over the different measurement occasions. Having a single measure has the advantage of facilitating interpretation and is very useful whenever two scales should be compared on their reliability. On the other hand, it is possible to obtain R_T values per time point, which can be important when one is interested in the evolution of reliability over the course of the study.

The R_Λ coefficient, even though structurally similar to R_T , bears a totally different interpretation. This measure expresses the reliability of the longitudinal sequence as a whole. It captures not the average reliability per time point, but the reliability of the information that is available when considering the repeated measures jointly. As a consequence, R_Λ will always increase when the number of measurements increases. Relevantly, this implies that we can always obtain a pre-specified level of reliability if the patient is followed long enough. Indeed, even if we only have to our disposal a scale that gives rise to a relatively large amount of measurement error, we can still increase the reliability of our conclusions by repeating the measurement over time.

The previous developments were based upon a longitudinal framework. However, in psychometric research much interest has gone to the study of reliability in the context of cross-sectional, multivariate measurement. We have illustrated that the same measures as proposed in the longitudinal context also apply when studying reliability in a multivariate setting. The R_T coefficient then expresses the average reliability per item whereas the R_Λ coefficient refers to the reliability of the information available in the entire scale.

Vangeneugden Tony. *Applying Psychometric Validation Methodology to Longitudinal Clinical Trial Data* (2008) – Promotor: Pr. G. Molenberghs.

Before questionnaires or measurement scales are used in clinical trials, the psychometric properties must be assessed to validate the measurement scale. Specifically, the reliability and the validity are evaluated. Reliability consists in determining the extent the measurement is free from random error. This can be performed through analyzing internal consistency and reproducibility of the questionnaire. The calculation of the intraclass correlation coefficient (ICC) is one of the most commonly used methods. The validity of a questionnaire is defined as the degree which the questionnaire measures to what it purports to measure. This can be performed through the analysis of content, construct and criterion validity. This psychometric validation is done in a separate and often in a rather small sample of stable subjects. However it is important to note that validity and reliability of a scale are not unique aspects of a scale but relative to the population in which the scale is used.

In this thesis we showed that these psychometric validation techniques can also be applied to longitudinal data collected in clinical trials. The goal is not to replace upfront psychometric validation, but rather to offer methodology to evaluate validity and reliability in the specific trial population at hand. This work provided a flexible framework to evaluate the actual performance of the scale in terms of reliability or validity. More specifically, to evaluate reliability, we used the Linear Mixed Model framework to develop a general formula to derive test re-test reliability in case of interval scaled data. The same framework was then used to extend reliability to generalizability testing. The purpose of this concept is to evaluate which factors influence reliability. Data from 5 clinical studies in schizophrenia were used to study test-retest reliability of the total Positive and Negative Syndrome Scale (PANSS). The same data was also used to evaluate generalizability of the total PANSS versus country and the baseline total negative PANSS subscore. Analysis showed the reliability of the total PANSS was not impacted by so much by the former but more by the latter factor.

This framework was then extended to binary case by means of the General Linear Mixed Model (GLMM) framework to derive approximate formulae for the ICC of reliability. These derivations allowed to derive reliability and generalizability of Clinical Global Impression (CGI) response captured in 4 clinical trials in schizophrenia. Additionally, the special case of count data was addressed. A closed form was derived to calculate the ICC of reliability and applied to number of seizures captured in a clinical study in epilepsy.

Finally, the GLMM framework was used to investigate correlation between joint longitudinal sequences of different measures. Similar a criterion validity, we evaluated the correlation between the PANSS total score and CGI response. Analyses showed that the correlation between the total PANSS and CGI response was as high as 0.75.

Van Sanden Suzy. *Statistical Methods for Microarray-based Analysis of Gene-expression, Classification, and Biomarker Validation* (2008) – Promotors: Pr. T. Burzykowski and Pr. Z. Shkedy.

One of the most important research questions that scientists still face today, is how do living organisms function, down to the cellular level. To find the answer, a substantial amount of research has been devoted to the study of the basic building block of life, DNA. However, the identification and localization of all of the approximately 20,000-25,000 genes in human DNA, a task which was completed in 2003, led to other questions: what is their function and how is their expression regulated? Gene expression is the process, by which the information carried by a gene is transformed into a protein. This process can be influenced by a number of internal (e.g., a disease) or external (e.g., the environment) factors.

Addressing these questions is, however, complicated by the fact that the genome of every living organism consists of a substantial number of genes. It is nearly impossible to examine expression levels gene by gene. The demand for techniques that would allow to simultaneously monitor a large numbers of genes was met by the development of DNA microarrays.

These technological advances in the field of genomics have brought about a new statistical research area, the analysis of data from high-throughput screening experiments. The main issue in such studies is the interest in a large number of parameters, while confronted with a small biological sample size. In addition, there is inherent “noise” in microarray data. The process of obtaining the gene-expression measurements is one of many stages. Each of these stages is vulnerable to the inclusion of unwanted systematic and random effects, possibly leading to bias in the results. Hence, there is a need for proper statistical procedures to deal with the above mentioned issues in the design and analysis of the experiments.

In the dissertation, we focus on data normalization, discovery of differentially expressed genes, and class prediction. A new transformation is proposed to pre-process data from different slides and to prepare them for further analysis. Several gene selection and classification methods are compared, with respect to their ability to classify samples to predefined groups, in an extensive simulation study. Furthermore, new modelling techniques are explored to discover differentially expressed genes. And in the final part, some methods are presented for biomarker detection.

University of Leuven (KULeuven)

Consentino Fabrizio. *Modelling the missingness: estimation, testing and model selection* (2009) – Promotor: Pr. G. Claeskens.

The term 'missing data' indicates the presence of missing observations in the data set of interest, which could be both in the response and the explanatory variables. When data are collected, in any possible scientific field, the occurrence of missing observations is quite probable; this presence has an impact on analyzing the data, since the standard statistical methods could fail in order to obtain reliable results. For this reason dealing with missing data has had an increasing impact in statistical analysis, leading, in the last years, to new techniques able to overcome these statistical problems, such as for estimation, model selection, etc.

The primary aim of the work presented in this dissertation is to develop new tools in order to analyze data sets with missing observations and to draw reliable conclusions from them. Nowadays many studies in different scientific fields have to deal with large amounts of information; these datasets include a large number of variables. Hence choosing the most relevant variables for the analysis is fundamental in order to obtain valid results. However problems could rise when missing data are present potentially leading to a failure of the procedures.

Since model selection could fail in the presence of missing data, a new variant of the classical Akaike information criterion (AIC) has been proposed. The new AIC is able to perform model selection when missing observations are present in the explanatory variables. This is done by exploiting the use of the EM algorithm and by modelling the covariates with missing observations. When the covariates contain missing observations, they are treated as random variables; hence modelling them has an impact on the model selection criteria with missing observations.

Using different distributions for this purpose could lead to better results. We propose a distribution selection criterion when either a normal or a t-distribution is chosen for modelling the missing covariates. In particular this is performed using a non iterative method, which is feasible when a logistic regression model is considered. Furthermore we consider estimation of parameters in these models.

Performing hypotheses tests is fundamental in statistics. The goal of the presented research in Chapter 4 is to develop nonparametric tests that are applicable to data sets with missing observations. The studied order selection tests do not require the specification of a particular parametric alternative hypothesis. The main advantage of these tests is that by considering orthogonal series expansions the tests have non-trivial power against a wide range of alternatives. For the construction of the tests a model selection criterion is used to select amongst the models in the series expansion. Our construction uses likelihood ratio tests based on multiple imputations in order to perform the order selection test in missing data situations.

Rizopoulos Dimitris. *Joint modelling of longitudinal and survival data* (2008) – Promotor: Pr. G. Verbeke

Many longitudinal studies collect information on outcomes such as time to infection or death, as well as covariates that vary with time. These covariates are usually measured intermittently, often at different times for each participant, and with substantial error. In such studies the prognostic value of these time-dependent covariates and/or the covariate process itself may be of interest, since it sheds light on the natural history of the disease. This type of studies have lead to a new and active area of biostatistical research that deals with the joint modelling of longitudinal and event time data.

In this thesis we investigate a number of issues in the joint modelling area. In particular, we theoretically investigate the effect of misspecifying the random effects distribution in parameter estimators and standard errors, especially as the number of repeated measurements per subject increases. Further, we propose an alternative parameterization for joint models using copulas for the random effects, and discuss sensitivity analysis issues. On the computational part of the thesis, a new type of Laplace approximation is developed that can efficiently handle multidimensional random effects vectors in joint models. Finally, we postulate a flexible model for the event outcome using B-splines under which the estimation of standard errors is facilitated.

Tsonaka Roula. *Models for handling coarsening and non-monotone missingness in clinical trials* (2008) – Promotors: Pr. G. Verbeke, Pr. E. Lesaffre and Pr. M. Hubert

In clinical trials often data are either missing or inaccurately recorded. This is the so called phenomenon of data coarsening that can occur in the form of missingness, censoring and grouping or rounding. For instance, consider a longitudinal study where a variable of interest (e.g., blood pressure) is repeatedly recorded in time on the same subjects. Even though data collection is often scheduled at pre-specified points in time, not all measurements are obtained. For various reasons the study participants may fail to appear at the study centers leading thus to incomplete response profiles. Moreover, in a study where the time to a particular event (e.g., death) is of interest, the true event times may not be recorded for all individuals. The end of the study may be reached or the subjects drop out from the study before experiencing the event. In this case the variable of interest is imposed to censoring. Furthermore, in survey studies the recorded variables are often prone to misreporting and grouping or rounding (e.g., age reporting for infants is typically done in weeks or months, for adolescents and adults it is truncated to the next lower year, etc.). Analysing such coarse responses as if they were complete can lead to incorrect inferences. Therefore the reasons of data coarsening need to be carefully considered and in some settings sophisticated methods of analysis are required.

In this thesis we have developed methods for handling data coarsening in clinical trials that occurs in the form of missingness and grouping or rounding. Specifically, in Chapter 1 we explain how coarseness arises in clinical trials and discuss the implications in analysis when it is not properly addressed. In Chapters 2 and 3 we concentrate on the analysis of non-monotone missing profiles, namely incomplete response profiles that arise when the subjects miss intermittently some of the scheduled visits. In particular, we consider the Shared Parameter Model (SPM) framework, in which a latent process (e.g., the true health status) known as random effects is assumed to affect both the longitudinal responses and the mechanism that produces the missing data. Two important issues are addressed; the choice of the random effects distribution and the interpretation of the model parameters. Regarding the random effects distribution, parametric assumptions are typically considered that can be unrealistic and affect the validity of inference. Therefore, in this thesis we leave this distribution completely unspecified. For the estimation of this model a semi-parametric maximum likelihood method is used. Regarding parameter interpretation, the use of random effects induces a conditional on the random effects interpretation, which may not always be desirable (e.g., when the population

treatment effect is required). Therefore a reparameterization is applied on the longitudinal model to allow for model parameters with a population averaged interpretation. This gives rise to the marginalized semi-parametric SPM presented in Chapter 3 for the analysis of incomplete longitudinal ordinal responses.

In Chapters 4 and 5 we deal with the analysis of grouped data, namely, responses that are a coarsened version of a latent variable of primary interest. In particular, the analysis of quality of life indexes is the main topic of this part of the thesis. Such indexes often follow non-standard distributions, namely J- or U-shaped, precluding classical parametric statistical approaches for analysis. Therefore, a new parametric approach, based on the logistic transformation, is proposed for analysing index data that can capture various shapes of distribution while allowing for covariate adjustments. The proposed model and the derived formulas for power and sample size calculations presented in Chapter 5 are applied to Barthel index evaluations of patients with an acute ischemic stroke.

Finally, in Chapter 6 we conclude with a discussion of the main findings of this doctoral research and refer to topics for further research.

Vermeulen Bart. *Design issues in conjoint analysis for market and non-market valuation* (2009) – Promotors: Pr. M. Vandebroek and Pr. P. Goos (UA)

The stated preferences of individuals for different products reveal a wealth of information to researchers and practitioners in market and non-market valuation. By exploring the preferences of individuals, one can gain insight in the economic values of the features of a product. In market valuation, the focus is on measuring the impact of new features on the individuals' purchases and on predicting market shares for new products. In non-market valuation, public goods are evaluated for which there exists no market and valuation is even more complex.

To explore the preferences of individuals, conjoint experiments have become a popular tool. In these experiments, a respondent is confronted with a number of hypothetical commodities or alternatives each described by its features or attributes. The respondent is requested to choose his or her preferred alternative, to rank or rate the alternatives or to choose the options he/she likes or dislikes most. This task is then repeated for a specified number of sets of hypothetical profiles. After analyzing these conjoint data by a random utility model, the resulting part-worth or utility coefficient estimates reflect the importance of the attributes for the individuals.

One of the key challenges in implementing a conjoint experiment is the statistical design of it. This involves the combination of attributes and their levels and the allocation of the resulting alternatives to the choice sets. Because the number of combinations of attributes and their levels can be huge even with relatively simple commodities, some theory is required to construct the most appropriate design. Depending on the objective of the analyst, an optimality criterion is used to do so. Often, the researcher is interested in maximizing the information on the utility coefficients. However, a researcher might put first another goal: e.g. measuring the individuals' willingness-to-pay to obtain a particular feature of a commodity. It is obvious that other optimality criteria will be required in each of these cases.

In this work, we discuss various topics related to the design of conjoint experiments. In the first chapter, we deal with conjoint choice experiments including a no-choice option which mimics an individual's market behaviour more realistically. We examine whether including this no-choice option in the design phase of the experiment improves the accuracy of the estimated utility coefficients and predictions of the individuals' choices.

In chapter 2 and 3, we consider *rank-order conjoint experiments*, in which the respondent is asked to rank all or a number of the alternatives in each choice set, and *best-worst choice experiment*, in which the respondent indicates the most and least preferred alternative in each choice set. We propose an optimality criterion for both types of experiments to develop optimal designs leading to precise estimates of the utility coefficients. We examine whether considerable improvements in terms of estimation and prediction accuracy are obtained by using the resulting tailor-made designs compared to benchmark designs which are often used in practice. Moreover, we measure the additional information of indicating a second choice in each choice set, i.e. choosing the second best or the least preferred option.

However, a conjoint experiment is not always focused on precise estimates of the utility coefficients. Measuring individuals' willingness-to-pay (WTP) is a frequently stated goal in non-market valuation. Obtaining accurate estimates of the willingness-to-pay necessitates an other design criterion than the one yielding precise estimates of the utility coefficients. Chapter 4 and 5 provide several design criteria to obtain precise willingness-to-pay estimates. We compare the designs constructed using the proposed criteria in terms of the precision of the WTP estimates.

Yu Jie. *Optimal design methodology for choice experiments in the presence of model uncertainty and consumer heterogeneity* (2009) – Promotors: Pr. M. Vandebroek and Pr. P. Goos (UA)

Conjoint analysis is by far the most preferred technique for exploring consumers' preferences for different features of an individual product or service. One of the most important strengths of this technique is the ability to develop market simulation models that can predict how consumers would react to product changes. This assists companies in determining what features a new product should have and how it should be priced.

In conjoint experiments, any product or service (called profile or alternative) is described in terms of a number of attributes. For example, a television may have attributes of screen size, screen format, brand, price, and so on. Each attribute consists of a number of levels. For instance, levels for screen format may be CRT, LCD, or Plasma. Respondents or test persons would be shown a set of products created from a combination of levels from all or some of the attributes and asked to rank or rate these products or choose their preferred one from this set of products. The last type of experiment is called the choice-based or conjoint choice or discrete choice experiment. It is popular because it imitates consumer behavior in real-life. It is a more realistic exercise for individuals to indicate which product they would purchase rather than rating or ranking since making choices is what they actually do in the marketplace. To date, discrete choice experiments have been extensively applied in many of the social sciences and applied sciences including marketing, product management, operations research, econometrics, transportation, environmental and health economics.

In a typical choice experiment, respondents need to indicate their preferred profile from each of several choice sets presented to them. As the number of combinations of attributes and levels increases, the number of potential profiles increases exponentially. This brings us to the question of what profiles should be used in the experiment and how to group them into choice sets such that the experiment can provide maximum information on the parameter estimates? The main focus of this thesis is to construct efficient experimental choice designs.

University of Liège (ULg)

Vanbelle Sophie. *Agreement between raters and groups of raters* (2009) – Promotors: Pr. A. Albert and Pr. G. Haesbroeck.

Agreement between raters on a categorical scale is a situation often encountered in practice. For example, one may want to test for agreement between two psychiatrists on the assessment of depression in out patients. We also may want to test for the equality of such obtained agreements with two different methods. The first part of this work concerns with the first situation and the second part with the second one.

In the first part of the work, the kappa-like family of agreement indexes is described in various situations: when agreement is searched between two singles raters or more, between an isolated rater and a group of raters and between two groups of raters. To quantify the agreement between two single raters, Cohen's kappa coefficient and intraclass kappa coefficient can be used for binary and nominal scales while weighted kappa coefficient concerns with ordinal scales. An interpretation of the quadratic and the linear weighting scheme is given. When agreement is searched between several raters, agreements indexes corresponding to Cohen's kappa and intraclass kappa coefficients are exposed. Then, the kappa-like family of agreement coefficients is extended to the case of an isolated rater and a group of raters and to the case of two groups of raters. The agreement coefficients are derived on a population model and reduce to the well-known Cohen's coefficient in case of two single raters. The proposed agreement indexes are also compared to existing methods, the consensus method and Schouten's agreement index. The superiority of the new methods on these methods is shown.

In the second part of the work, methods to compare several agreement indexes are presented. Firstly, the method proposed by Fleiss (1981) to compare several independent agreement indexes is developed. Then, a bootstrap method is presented to compare two dependent agreement indexes. This method was extended by to compare more than two dependent agreement indexes. The methods cited above can be applied to compare all kind of agreement indexes introduced in the first part of the work. Finally, regression methods to test the effect of continuous and categorical covariates on the agreement between two or several raters are presented. This includes the weighted least square method, allowing only for categorical covariates and a regression method based on two sets of generalized estimating equations. The latter method was developed for the intraclass kappa coefficient, Cohen's kappa coefficient and the weighted

kappa coefficient. Finally, an heuristic method, limited to the case of independent observations is presented which turns out to be equivalent to the generalized estimating equation approach. The regression methods have still to be adapted to allow for agreement indexes between a single and a group of raters and between two groups of raters.

University of Louvain-la-neuve (UCL)

Motta Giovanni. *Evolutionary factor analysis* (2009) – Promoters: R. von Sachs and C. Hafner

Linear factor models have attracted considerable interest over recent years especially in the econometrics literature. The intuitively appealing idea to explain a panel of economic variables by a few common factors is one of the reasons for their popularity. From a statistical viewpoint, the need to reduce the cross-section dimension to a much smaller factor space dimension is obvious considering the large data sets available in economics and finance.

One of the characteristics of the traditional factor model is that the process is stationary in the time dimension. This appears restrictive, given the fact that over long time periods it is unlikely that e.g. factor loadings remain constant. For example, in the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), typical empirical results show that factor loadings are time-varying, which in the CAPM is caused by time-varying second moments.

In this thesis we generalize the tools of factor analysis for the study of stochastic processes whose behavior evolves over time. In particular, we introduce a new class of factor models with loadings that are allowed to be smooth functions of time. To estimate the resulting non-stationary factor model we generalize the properties of the principal components technique to the time-varying framework. We mainly consider separately two classes of Evolutionary Factor Models: Evolutionary Static Factor Models (Chapter 2) and Evolutionary Dynamic Factor Models (Chapter 3).

In Chapter 2 we propose a new approximate factor model where the common components are static but non-stationary. The non-stationarity is introduced by the time-varying factor loadings, that are estimated by the eigenvectors of a non-parametrically estimated covariance matrix. Under simultaneous asymptotics (cross-section and time dimension go to infinity simultaneously), we give conditions for consistency of our estimators of the time varying covariance matrix, the loadings and the factors. This paper generalizes to the locally stationary case the results given by Bai (2003) in the stationary framework. A simulation study illustrates the performance of these estimators. The estimators proposed in Chapter 2 are based on a nonparametric estimator of the covariance matrix whose entries are computed with the same smoothing parameter. This approach has the advantage of guaranteeing a positive definite estimator but it does not adapt

to the different degree of smoothness of the different entries of the covariance matrix.

In Chapter 5 we give an additional theoretical result which explains how to construct a positive definite estimate of the covariance matrix while while permitting different smoothing parameters. This estimator is based on the Cholesky decomposition of a pre-estimator of the covariance matrix.

In Chapter 3 we introduce the dynamics in our modeling. This model generalizes the dynamic (but stationary) factor model of Forni et al. (2000), as well as the non-stationary (but static) factor model of Chapter 2. In the stationary (dynamic) case, Forni et al. (2000) show that the common components are estimated by the eigenvectors of a consistent estimator of the spectral density matrix, which is a matrix depending only on the frequency. In the evolutionary framework the dynamics of the model is explained by a time-varying spectral density matrix. This operator is a function of time as well as of the frequency. In this chapter we show that the common components of a locally stationary dynamic factor model can be estimated consistently by the eigenvectors of a consistent estimator of the time-varying spectral density matrix.

In Chapter 4 we apply our theoretical results to real data and compare the performance of our approach with that based on standard techniques. Chapter 6 concludes and mentions the main questions for future research.

Teodorescu Bianca. *General conditional linear models with time-dependent coefficients under censoring and truncation* (2008) – Promotor: Pr. Ingrid Van Keilegom

In survival analysis interest often lies in the relationship between the survival function and a certain number of covariates. It usually happens that for some individuals we cannot observe the event of interest, due to the presence of right censoring and/or left truncation. A typical example is given by a retrospective medical study, in which one is interested in the time interval between birth and death due to a certain disease. Patients who die of the disease at early age will rarely have entered the study before death and are therefore left truncated. On the other hand, for patients who are alive at the end of the study, only a lower bound of the true survival time is known and these patients are hence right censored.

In the case of censored and/or truncated responses, lots of models exist in the literature that describe the relationship between the survival function and the covariates (proportional hazards model or Cox model, log-logistic model, accelerated failure time model, additive risks model, etc.). In these models, the regression coefficients are usually supposed to be constant over time. In practice, the structure of the data might however be more complex, and it might therefore be better to consider coefficients that can vary over time. In the previous examples, certain covariates (e.g. age at diagnosis, type of surgery, extension of tumor, etc.) can have a relatively high impact on early age survival, but a lower influence at higher age. This motivated a number of authors to extend the Cox model to allow for time-dependent coefficients or consider other type of time-dependent coefficients models like the additive hazards model. In practice it is of great use to have at hand a method to check the validity of the above mentioned models.

First we consider a very general model, which includes as special cases the above mentioned models (Cox model, additive model, log-logistic model, linear transformation models, etc.) with time-dependent coefficients and study the parameter estimation by means of a least squares approach. The response is allowed to be subject to right censoring and/or left truncation.

Secondly we propose an omnibus goodness-of-fit test that will test if the general time-dependent model considered above fits the data. A bootstrap version, to approximate the critical values of the test is also proposed.

In this dissertation, for each proposed method, the finite sample performance is evaluated in a simulation study and then applied to a real data set.