LIKELIHOOD-BASED ESTIMATION METHODS FOR CREDIT RATING STOCHASTIC FACTOR MODEL

MAYGOL BANDEHALI

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN ECONOMICS
YORK UNIVERSITY
TORONTO, ONTARIO

JULY, 2020

Abstract

After recent financial crises, finding the right methodology to control and regulate risks that arise in the financial markets became the most important task of a bank. A bank faces various risks, among which the most important risk any bank encounters is the credit risk. In the context of Basel II, the major components of credit risk management are credit rating migration matrices. The analysis and modelling of the dynamics of these matrices are crucial for predicting the future risk of borrower's default, the regulatory capital requirement, credit approval, risk management and internal credit allocation. This thesis is an empirical investigation of various estimation methods for the analysis of the dynamics of credit rating matrices. More specifically, the thesis discusses the statistical estimation of the latent factor ordered-Probit model, which is also known as the stochastic credit migration model, a homogeneous nonlinear dynamic panel model with a common unobserved factor, to determine the dynamics of credit ratings transition probabilities. In particular, this thesis presents three maximum likelihood estimation methods of the latent factor ordered-Probit model. The first two methods rely on analytical approximation of the true log-likelihood function of the latent factor ordered-Probit model based on the granularity theory and are discussed in Chapter two. The third method is maximum composite likelihood estimation of the latent factor ordered-Probit model which is the new approach to estimate the latent factor ordered-Probit model and is discussed in Chapter three.

Chapter 1 provides the literature review on the dynamics, estimation and modelling the credit rating transition matrix. It describes the credit rating dynamics and explain how individual credit rating histories are used to predict the future risk of a given borrower or

set of borrowers. The general framework of the ordered qualitative models that are used to model the credit rating dynamics is explained. The notation and general assumptions of a latent factor ordered-Probit model are introduced and the statistical inference of the latent factor ordered-Probit model is discussed.

Chapter 2 reviews two maximum likelihood estimation methods of the latent factor ordered-Probit model which are the two-step estimation and joint optimization. The methods considered rely on analytical approximations of the log-likelihood function based on the granularity theory. The effect of the underlying state of the economy on corporate credit ratings is inferred from the common factor path. The estimated model allows us to examine the effects of shocks to the economy, i.e. the stress testing at the overall portfolio level, which is also a required element of the execution of Basel II. The stress scenarios are selected to evaluate the stressed migration probabilities and relate them with the state of the economy. The empirical results are obtained from the series of transition probabilities matrices provided by the internal rating system of Credit Agricole S.A. bank over the period 2007 to 2015.

Chapter 3 introduces the maximum composite likelihood estimation method for the latent factor ordered-Probit model for credit migration matrices. The computational complexity of the full information maximum likelihood in application to the stochastic migration model is the main motivation to introduce a new method, which is computationally easier and provides consistent estimators. The new method is illustrated in a simulation study that confirms good performance of the maximum composite likelihood estimation.

Acknowledgements

I am tremendously grateful to my supervisor, Professor Joann Jasiak, for her invaluable supervision, her endless guidance, and support. It was a privilege for me to be her student. I would like to express my deepest gratitude to Prof. Christian Gourieroux, for his willingness contribution, feedback, and collaboration to shape this thesis.

I want to thank all of my committee members, Prof. Augustine Wong and Dr. Antoine Djogbenou in which their constructive feedback, comments, and great suggestions helped to improve this dissertation. My special thanks to Dr. Antoine Djogbenou, for taking the time and effort to guide me through many conceptual details of my thesis, and providing me with excellent advice, and insightful suggestions.

Eventually, I will forever be grateful to my parents for their continuous love, support, and encouragement during all steps of my life including my graduate studies. I am extremely thankful to my siblings Maziar and Sogol, for their conditions less support, and being a source of love and energy to complete this dissertation. My sincere appreciation to Dr. Amirhossein Ahadi, for his insightful comments, and professional consults during my Ph.D. studies.

Finally, I wish to thank all the people whose assistance was a milestone in the completion of this thesis, and without their help and contributions, this thesis would not have been fulfilled. I dedicate this thesis to my family, for their constant support and unconditional love.

Contents

List of Tables

List of Figures

Introduction

One of the most active areas of the recent financial research is credit risk, which is the most important risk confronted by a bank in its traditional intermediation. This is the risk that an obligor defaults and does not honour its responsibility to service debt. The fundamental instrument of any credit risk analysis is the credit rating, which is the key measure to quantify credit risk. There are two kinds of credit ratings: one is routinely provided and updated by the rating agencies, such as the Moody's, Standard & Poor's and Fitch, and called external ratings. The other is measured and updated by the internal scoring systems of banks, and known as internal or shadow ratings. The credit ratings concern bond issuers such as corporations, individuals, retail and countries. The rating histories are valuable tools for the prediction of the upcoming risk of a specific borrower, or a set of borrowers being downgraded or defaulted. The ratings change over time. Past changes in credit quality of an obligor are documented by credit migration or transition matrices which track the movement of firms across rating categories over a given period of time and within a given group of bond issuers. The record of changes in the credit rating of an obligor is a key component of a bank's credit scoring. The past scores allow banks to evaluate and price credit risk more accurately. In the literature, there exists few rating migration studies focused on internal rating systems of financial institutions, due to the confidentiality of data. Credit migration or transition matrices have become an essential element of many risk management methods including the risk assessment, from the VaR on corporate credit portfolios (CreditVar), to bond pricing models [see, e.g. Jarrow and Turnbull (1995), Jarrow et al. (1997)], to credit derivative pricing models [see, e.g. Kijima and Komoribayashi (1998), Acharya and R.K.

Sundaram (2002)] and banks' credit capital determination.

This thesis focuses on an important branch of research which is the importance of the internal credit rating and credit scoring economic consequences, particularly after the last financial crisis in 2008. The financial crisis of 2007-2008 is the most severe financial crisis since the Great Depression in the 1930s. It was a series of crisis that undulated through the financial system and finally, the economy. It started in US subprime mortgage market and led to a severe global financial meltdown. Many experts failed to identify the risk of subprime mortgages and other risky loans. The key enablers of the financial meltdown were the three credit rating agencies (Standard & Poor's, Fitch, and Moody's). Large financial institutions trusted the credit ratings of these agencies. These agencies rated several collateralized debt obligations as "AAA" (prime), while most of them were not even prime investments. Nearly 45,000 mortgage-related securities were rated as "AAA" by Moody's from 2000 to 2007. A large number of them defaulted or lost vigorously in value and led the financial institutions to lose tremendous amounts of capital. These agencies have been accused of overstated appraisals of risky loans, giving financial specialists bogus certainty that they were of for contributing. During the financial crisis, the role of the credit rating agencies remains highly criticized and mostly unaccountable. Many policymakers, multilateral agencies, academics and investors were largely surprised by the crisis. Financial policymakers and decision makers reacted to the near-meltdown of the global financial sector on a real-time basis, through a dynamic set of monetary and fiscal policy measures to curb the crisis. The Basel Committee on Banking Supervision significantly reduced the banks' dependence on external ratings during the calculation of capital requirements in 2014. This was done to force banks to improve their own risk assessments of bad loans. The position of credit rating agencies was not limited solely to the US. According to the European Central Bank (ECB), credit rating agencies' inaccurate sovereign ratings led the worsening of the euro-zone debt crisis. Following the sovereign downgrades, the EU proposed that the formation of a European credit rating agency would question the oligopoly of the three major rating agencies in Europe. It is also not shocking that the magnitude of this economic crisis has been underestimated for most of 2008. The emergence of Basel III is an essential part of the regulator's reaction to the financial market crises and to the endogenous underestimation of systematic risk by major players.

In recent years, the requirements regarding the operation of an internal rating-based (IRB) approach, the associated internal credit rating studies, and the corresponding rating migration analysis have followed in the framework of the Basel II and III agreements. The Basel Committee on Banking Supervision (BCBS), was formed in 1974. The aim of the committee is to improve the financial stability of banks by introducing the Basel Accords. The Basel Accords contain three sets of banking regulations (Basel I, II and III) which concern three major risks, the banks are facing such as, the capital risk, the market risk, and the operational risk, to cover unexpected losses of banks so that they can meet their obligations. In Basel I, the capital needed for supporting different assets is measured by a standardized risk-weighting approach while under Basel II and III, the committee allows banks to use their own internal risk models. Under the internal ratings-based approach (IRB), banks use their own internal credit ratings to predict the probability of default (PD), the loss given default (LGD), the exposure at default (EAD) and to estimate the entire matrix of credit transition probabilities. The rating histories are used for different purposes, such as credit score approximation and assessment of risk on a credit portfolio of a large number of borrowers. However, the ratings are serially correlated, which means that the lagged ratings have an effect on the current ratings. Moreover, there is cross-sectional correlation which generally concerns migration correlation or more precisely default correlation among the obligors.

Motivated by this observations, it is important to consider how rating systems are conceptualised, developed, implemented and utilised in risk management to recognise how banks conduct their business lending role and how they want to monitor risk exposures. Generally, banks and rating agencies gather amounts of public and private financial data while producing credit ratings and then produce regularly revised credit ratings based on a mixture of standardized and empirical econometrics models Field specialists judgments. Previous research on automated credit rating generation at the company and industry level in the finance and economics community ranged from discriminatory analysis Pinches and Mingo (1973), logistical regression Ederington (1985), decision analysis, and artificial intelligence Dimitras et al. (1996). recent years, multiple strategies in data mining, including computational and machine-learning enterprise credit risk assessment techniques have been applied. [see, e.g. Angelini et al. (2008); Khashman (2010)]. Such experiments have shown that neural networks outperform conventional mathematical and econometric approaches in the problems of credit rating prediction. Although, this research is not about the machine-learning approach but uses the statistical learning tools, such as the principal component analysis to obtain the path and the number of factors that drives the credit rating migrations. Thus establishing the proper methodology for estimating migration matrices and forecasting the probability of default and rating transition are important. These methods allow banks to evaluate and price credit risk more accurately, as well as help them to improve the assessment of minimum capital requirements.

This thesis provides an empirical examination of various estimation methods for the analysis of the dynamics of credit rating matrices. More specifically, the credit rating migrations are modeled by the factor ordered-Probit model which is also known as the stochastic credit migration model. The latent factor ordered-Probit model is a homogeneous nonlinear dynamic panel model with a common unobserved factor representing a common systematic risk. It allows us to determine the dynamics and forecast of the credit ratings transition probabilities. This thesis presents three maximum likelihood estimation methods of the latent factor ordered-Probit model. The first two methods are the two-step estimation and the joint optimization. These methods rely on analytical approximation of the true log-likelihood function of the latent factor ordered-Probit model based on the granularity theory. The third method is maximum composite likelihood estimation of the latent factor ordered-Probit model. The maximum composite likelihood method is a new approach to estimate the stochastic factor ordered-Probit model. Moreover, this thesis provides empirical results for the first two estimation methods, the two-step and joint optimization, which are obtained from the series of transition probabilities matrices provided by the internal rating system of Credit Agricole S.A. bank over the period 2007 to 2015. A simulation study is provided as an illustration of the new maximum composite likelihood estimation approach.

This thesis makes the following contributions within the general framework and discusses them in chapters 2 and 3:

(A) *an application of stress testing of latent factor ordered-Probit model through shocks applied to the factor*

(B) *an interpretation of latent factor estimated from the latent factor ordered-Probit model from data in internal ratings of a French bank*

(C) *a contribution to macroeconomic stress-testing via migration matrices based on an internal rating system*

(D) *an empirical comparison of the two-step estimation and joint optimization approaches for the latent factor ordered-Probit Model*

(E) *a derivation of introductory results for identification, consistency and asymptotic normality of the maximum composite likelihood estimator of the latent factor ordered-Probit Model*

(F) *a simulation study of the new maximum composite likelihood estimator of the latent factor ordered-Probit Model*

The first chapter, presents the literature review on the dynamics, estimation and modelling of credit rating transition matrices. The comprehensive review of the literature that uses these matrices to model and predict future risk in a large credit portfolio is discussed. First, we describe the credit rating dynamics and explain how individual credit rating histories are used to predict the future risk of a given borrower or set of borrowers. Then, we explain the general framework of the ordered qualitative models that are used to model the credit rating dynamics. More specifically, in order to model and predict the credit rating matrices, we employ a factor-Probit model with stochastic credit rating migration matrices which are driven by a latent factor. The effect of the business cycle on corporate credit ratings can be inferred from the common latent factor path. We introduce the notation and general assumptions of a latent factor ordered-Probit model. Next we discuss the statistical inference of the latent factor ordered-Probit model. Due to computational complexity of the true log-likelihood function in latent factor ordered-Probit model, this thesis examines approximation-based estimation methods which simplifies the true log-likelihood for the latent factor ordered-Probit model and offer fast and reliable estimation. These approximation-based estimation methods are discussed in the next two chapters of this thesis.

In the second chapter, we compare two estimation methods for the latent factor ordered-Probit model in an application. The methods considered maximized the analytical approximations of the log-likelihood function based on the granularity theory. The computational complexity of the full information maximum likelihood in nonlinear dynamic panel models such as this stochastic migration model is the main motivation to use the approximation-based estimation methods. When compared to the computationally demanding simulation methods, the analytical approximations of the true likelihood function offer fast and reliable estimation that provides the output quickly and satisfy all the regulatory conditions. The advantage of these methods is mainly computational ease. They provide consistent estimators. The empirical results are obtained from the series of transition probabilities matrices provided by the internal rating system of Credit Agricole S.A. bank in France over the period 2007 to 2015. The stochastic factor ordered-Probit framework is used to model, predict the credit rating probabilities consistently with the Basel II implementation outline. The estimated model allows us to examine the effects of shocks to the economy, i.e. for the stress testing at the overall portfolio level which is also a required element of the execution of Basel II. The stress scenarios are selected to evaluate the stressed migration probabilities and relate them with the state of the economy.

Chapter 3 introduces the new maximum composite likelihood estimation method for the latent factor ordered-Probit model for credit migration matrices. We provide introductory results on the consistency of the estimators and their asymptotic normality. The new method is illustrated in a simulation study that confirms good performance of the maximum composite likelihood estimation approach.

Chapter 1

Review of Literature on Credit Rating Transition Matrices

1.1 Introduction

In the financial industry, finding appropriate methodologies to control and regulate risks has always been in the center of attention, especially after the last financial crisis of 2008. Credit hazard measurement stays a critical area of pinnacle precedence in banking finance immediately implicated inside the current global financial crisis. By the Basel Committee on Banking Supervision (2000) terminology, credit risk is defined as "the potential that a bank borrower or counterparty will fail to satisfy its obligations according with agreed terms". According to Basel II and III guidelines, banks are allowed to include their very own estimates of different threat exposures, adopting inner threat rating systems, with the intention to realistically verify their regulatory capital requirements, referred to as the "internal-ratings-based" (IRB) approach. The fundamental instrument of any credit risk analysis is the credit rating, which is the key measure to quantify credit risk.

In the recent years, the Basel III regulation has revealed the importance of credit rating transition matrices for the purpose of credit risk modelling in the financial industry. Modelling, estimation and forecasting of credit rating changes became more and more important in modern risk management. The first step to assess the credit exposure and potential losses is to estimate the migration probabilities and probabilities of default which are the basic inputs in the analysis of the systematic risk. The second step is to model the ratings dynamics and forecasting these matrices. There is an ongoing debate regarding which model is better for the purpose of forecasting default, rating migrations and/or hedging performance. The objectives of this literature are as follows:

- To find an appropriate methodology for estimating the credit transition matrices in order to provide a precise evaluation of the capital requirements.
- To assess the prediction power of models for probability of default and migration matrices and to evaluate various stress scenarios.

In terms of the methodology for estimating the credit transition matrices, there exists three multi-state Markov chain methods for estimating the transition matrices. The first one is the "Cohort" approach which is well-known in the industry. The second one is the "Duration" methodology which is very popular in academic research. The third method is the "Mixed Time Duration" which is based on the survival analysis. These methods have different time settings and data observation patterns. In the "Cohort" method, the matrix of transition probabilities is based on a discrete-time homogeneous Markov chain. The "Duration" methodology incorporates the possibility of successive downgrades leading to default in the continuous time set up. The "Mixed Time Duration" considers a mixing of discrete and continuous time observations. The underlying assumptions of first-order Markov and time homogeneity of the transition matrix for these three methods may be hard to maintain over a long horizon. Thus, few practical alternatives are offered in the literature [see, Bangia et al. (2002); Mählmann (2006)]. As of the modelling transition matrices, there are two major techniques for modelling and forecasting ratings migrations. The first one is the factor model and the second major techniques involves numerical adjustment methods based on macroeconomic situation. The factor model consists of two main frameworks, structural or reduced form. The factor models can be used to build an unobserved credit indicator to model the migration matrix in different frameworks [see, Kim (1999); Wei (2003);

Bae and Kulperger (2009); Berteloot et al. (2013)]. The numerical adjustment methods are convenient and are used by financial institution [see, Jarrow et al. (1997); Lando (2000)]. However, their accuracy has been questioned by Trück (2008).

Chapter 1 is organized as follows. In Section 1.2, the literature on estimating, modelling and forecasting the credit ratings migration matrices is discussed. Section 1.3 reviews the literature on the estimation methods of the latent factor ordered-Probit model. The specification of credit rating dynamics is described, and the notation and some general assumptions are explained in Section 1.4. In Section 1.5, the general framework of the ordered qualitative model and the latent factor ordered-Probit model are explained. Section 1.6 explains the estimation methodology for the latent factor ordered-Probit model. Section 1.7 concludes.

1.2 Review of Literature on Estimation and Modelling of Credit Ratings Migration Matrices

1.2.1 Estimation of Credit Migration Matrices

The first experimental credit transition matrix was introduced by Jarrow et al. (1997). They developed a discrete time homogeneous Markov model for the term structure of credit risk spreads. They used the Cohort methodology to construct the transition matrices. Each element in the matrix, represents the probability of transition from one rating category to another. It is calculated by dividing the number of firms that moved from one state to another by the total number of firms in the initial rating category. However, the assumption of discrete time in this context, has some drawbacks which make the Cohort methodology inefficient. One of the drawback is that the estimation of default probabilities of corporations with high rating grades is not accurate due to the scarcity of data on large corporation with high rating category. The other drawback is that the default through a sequence of downgrades or indirect defaults is ignored in the cohort methodology.

The continuous time approach (i.e., the Duration method) seems to provide a potential solution to the problems of the Cohort methodology. The aim of the continuous time framework is to capture the possibility of sequential downgrades of an obligor from higher to lower ratings until defaults occur. Two continuous time estimators for credit rating matrices are provided by Lando and Skødeberg (2002). These estimators differ in term of assumptions on the migration intensities. One of the estimator, called the Aalen-Johanson estimator, which allows transition intensities to be time varying over the business cycle while the other estimators assume that transition intensities are time invariant and so estimates a so-called infinitesimal generator matrix in continuous time Markov chains. Most of this literature considers continuously observed transition matrices data which are constructed by the rating agencies since the internal ratings are not continuously monitored, which could be a concern for the Duration methodology.

In the literature on the survival analysis method, Mählmann (2006) proposes a solution to the issue by a maximum likelihood procedure derived from the earlier work of Kay (1986). The key idea in the maximum likelihood procedure is to account for the interval-censoring time issue and the censored states issue to treat the likelihood function in a continuous-time setting. The interval-censoring time issue refers to the fact that the exact transition time is rare to witness, but could be know to occur within a particular interval. In order to address this issue, the maximum likelihood introduced by Mählmann (2006) uses different likelihood function from the Duration method. Moreover, the censored states refers to situation that an obligor does not default within the period studied and is not present at the end of study, it is in a censored state at the end of the period studied. Mählmann (2006) used data from a German bank, and compared the industry standard Cohort method and the maximum likelihood procedure. He showed that the cohort method overestimates default risk compared to the maximum likelihood approach. He shows that both the Duration and Cohort approaches are not adequate for estimating the transition matrices of internal rating data. The maximum likelihood procedure is actually the homogeneous multi state Markov model for mixed discrete-continuous time observations data in the literature of survival analysis. A clear description of a Homogeneous multi state Markov model is provided by

Commenges (2002). His data set contains clinical trials observations which are hard to document in a continuous time framework. The process of credit rating evaluation and clinical trials are very similar in the sense that both processes include the treatment and response. The characteristics of data in both processes are time-to-event and transitions between states and both have default and death states which are absorbing states. The internal rating systems should use the Kay's maximum likelihood estimators or Mixed-Time Duration since, they include mix discrete-continuous observation data.

Another challenge in recent literature is the dilemma whether a rating system is point-in-time (PIT) or through-the-cycle (TTC) which is an important classification of rating systems. In January 2001, an official definition of difference between the PIT and the TTC credit ratings was provided by the Basel Committee on Banking Supervision. Basel Committee obviously believes in the PIT approach, and recommends default risk measures over a short horizon. For the default risk metrics computation, all relevant information and all variables are supposed to be used in the process. The TTC approach considers longer horizons, of five or more years. Despite many advantages of the PIT rating system, less volatility and smooth transitions over cycles are obtained from the TTC rating system. These two approaches are considered as opposite types of possible rating procedures. Most of external rating agencies are using the TTC approach while the internal rating procedure most likely are combinations of these two approaches.

1.2.2 Modelling of Credit Migration Matrices

The financial institutions need to model and predict the credit transition matrices, in order to analyse, manage and assess the risk linked to a credit portfolio. Nowadays, there exists several risk management tools such as J.P. Morgan's Credit Metrics and McKinsey's Credit Portfolio View which are based on the estimation of transition probabilities. The definition of a credit event which indicates that the loss has occurred is one of the most crucial issues in the credit risk modelling. Generally credit risk models can be divided into two broad categories: the structural models and the intensity-based models or reduced-form models.

Black and M.Scholes (1973) and Merton (1974) introduced the structural models. In the structural models, default can be explained by a specific trigger point. For example, it can be caused by a decrease in asset value below some thresholds (i.e. the value of debt). Jarrow and Turnbull (1992, 1995), Duffie and Singleton (1999) introduced the reduced-form models. These models assume that a default intensity causes the defaults and not a specific trigger event. However, the default intensity might depend on changes in external factors such as GDP growth, inflation, unemployment and interest rates.

In the recent literature, there exists many studies that explore a variety of credit rating histories. Most of these studies focused on investigating the relations between ratings dynamics and the business cycle. Essentially, these studies are focused on the reasons of joint rating movements, which are often connected to the business cycle patterns. Many authors consider some observable variables as proxies for the stage of the business cycle, such as the risk-free interest rate [see, e.g. Nickell et al. (2000), Bangia et al. (2002)] or the unemployment rate [see, e.g. Rösch (2005)]. The models with an observable factor are not suitable for predicting future risk as the factor itself needs to be predicted. Several articles in the literature studied the default correlation to analyse joint movement of individual risks and considered unobserved factors. One of the advantages of this type of factor is that it helps predict future risks on credit portfolio easier than an observable factor. The history of ratings of a firm, and of joint ratings of multiple firms can help forecast the future portfolio's risk. Schönbucher (2000), Gordy and Heitfield (2001) and Gordy and Heitfield (2002) examine this type of model. Gagliardini and Gouriéroux (2005) develop a model to explain the rating transitions with unobservable time varying stochastic factors. Moreover, other studies such as Koopman et al. (2008), Creal et al. (2013) and Creal et al. (2014) also considered latent factors in their model to predict future risks on credit portfolios.

Banks are required by the Basel Committee on Banking Supervision (2001, 2003), to conduct stress tests on their credit portfolios. The stress tests provide a comprehensive assessment of banks ability to survive unexpected market changes. Stress testing became a focus of interest in risk management since the melt-down of the U.S. sub-prime mortgage market. In many countries, it is mandatory to conduct stress tests for credit risk management. Under Basel II credit risk framework (regulation approach), the credit rating migration matrices are essential to calculate the marked-to-market (MTM) losses and related capital requirements. Given the regulation suggested by the Basel committee, the literature on credit risk can be divided into three main categories:

1. The first category refers to studies that introduce individual covariates for determining a score function and predicting risk. The interest in these studies is in defining the scoring system. Chava and Jarrow (2002), Bharath and Shumway (2004), and Duffie and Wang (2004) are examples of these studies. In the latter study, the score function is based on the firm size, distance-to-default, and the firm's earnings. Thus firm's characteristics are introduce into their model, in the absence of microeconomic covariates.

2. The second category are studies which include observable macro-variables in the model and analyse the joint movements of individual risks. Understanding the joint rating movement is the focus of this literature. For instance, Nickell et al. (2000), measure the dependence of rating transition probabilities on the industry and the residence of the obligor, and on the stage of the business cycle. The specific connection between the macroeconomic volatility and asset quality is provided by credit migration matrices in Bangia et al. (2002). They divides the economy into two states, recession and expansion, and shows that there is a substantial difference between the loss distribution of credit portfolio in these two states. These two articles use external rating data. As for the internal rating system, Krüger et al. (2005) uses data from the Deutsche Bundesbank and shows that rating transitions depend on macroeconomic factors. Also, Mählmann (2006) uses the internal rating data of a German bank to show the relationship between rating migrations and the business cycle.

3. The third category refers to studies which include unobservable factors in the model and analyse the joint movement of individual risks. The main topic is default correlation

rather than risk dynamics. Generally, the current unobservable factors models are not suitable for prediction purposes, due to the assumption of serial independence. [see, e.g. Schönbucher (2000), and Gordy and Heitfield (2001, 2002)]. Two time independent factors are introduced in Duffie and Wang (2004). One of these factors is personal income growth which is observable and the other one is an unobservable factor. Most of these models do not have the capacity of forecasting ratings migration. Gagliardini and Gouriéroux (2005) introduce serially dependent unobservable factors which allow the future risk in a large credit portfolio to be predictable. Their model extends the standard and basic stochastic intensity model introduced by Cox and Lewis (1972) by considering a set of Markovian processes with stochastic transition matrices. Lando (1998), Duffie and Singleton (1999), Duffie and Lando (2001) and Gagliardini and Gouriéroux (2005) are other examples of studies which consider serially dependent unobservable factors which are more appropriate for prediction purposes.

1.2.3 Factor Ordered-Probit Model

In order to forecast rating migrations, two techniques are used in the literature. The first class of approach uses factor models to construct an unobserved credit indicator under different (structural or intensity-based) frameworks. A one-factor model is used to adjust migration matrices to a shift in the business cycle index. Belkin et al. (1998) and Kim (1999) include business cycle in their model and use a one-factor model to forecast credit rating migration. Wei (2003) introduces two factors in this model and one of the factor is the initial rating of a borrower. Another two-factor model is developed by Figlewski et al. (2012), which is based on intensity-based model. Figlewski et al. (2012) incorporate two factors that are a firm-specific and a macroeconomic variable in their model. They find that ratings-related factor and the macroeconomic factors have a strong effect on credit ratings migration. Bae and Kulperger (2009) and Berteloot et al. (2013) use an ordered logistic regression model with exogenous and endogenous covariates to predict transition matrices.

Another major technique for modelling migration matrices involves numerical adjustment methods [see, Jarrow et al. (1997); Lando (2000)]. This class of models, provides numerical adjustment methods to link historical transition matrices and observed market bond prices to obtain risk neutral migration matrices being in line with market credit spreads. Note that Lando (2000) gives some extensions of the adjustment methods that will also be included in the analysis. Although, this technique is convenient and is used by financial institutions, its accuracy has been questioned by Trück (2008). He compared these two approaches, and finds that one-factor model performs better in-sample and out of sample for predicting transition matrices.

Generally, the ordered polytomous model is the specification that the rating agencies and the Basel committee are considering to model the dynamics of credit rating transitions. In credit risk analysis, the structural and intensity-based models are special cases of ordered polytomous models with different assumptions on the error terms [Gagliardini and Gouriéroux (2005)]. Many articles use the factor model, such as Bhatia et al. (1997), Crouhy et al. (2000), Gordy and Heitfield (2001), and Bangia et al. (2002). Usually the credit risk literature use the Probit, Logit and Gompit models which are different types of the polytomous model. A Probit model is based on the assumption of normally distributed error terms and is refereed to as the structural model. Kim (1999), Nickell et al. (2000) and Trück (2008) are examples of studies which use the Probit model. In a Gompit model, the error terms have the Gompertz distribution. It is usually refereed to as the reduced form model. For instance, Figlewski et al. (2012) use the Gompit model. When the error terms have a logistic distribution, the model is called the Logit model and this model often is used in the industry. For example, Bae and Kulperger (2009) and Berteloot et al. (2013) use the Logit model for modelling of probability of default.

The discrete, ordinal nature of the credit ratings and rating transitions can be modelled by the ordered-Probit model. Nickell et al. (2001) employ the ordered-Probit model in the study of the stability of credit rating migrations. They find strong effects of the business cycle, and various industrial and domicile factors on the rating transitions. [1]

[1] Bliss (1934) introduced Probit model. It is usually estimated by the maximum likelihood which was proposed by Ronald Fisher in 1935.

This chapter follows the literature which includes unobservable factors in the model used for the analysis of the joint movements of individual risks. To do so, a vector Markov processes with serially dependent stochastic transition matrices will be considered. To model the dynamic of rating transitions, the ordered-Probit model will be used. The analysis is carried out in the framework of an ordered qualitative variable model [see, e.g Basel Committee on Banking Supervision (2000); Cheung (1996); and Nickell et al. (2000)]. This method explicitly permits for discreteness of rankings and accommodates naturally the ordering from low to high credit quality. Generally, the maximum likelihood is the estimation method for the ordered-Probit model. However, in our case, the maximum likelihood approach is difficult due to computational complexity. The ordered-Probit model includes unobservable (latent) variables. Therefore, large multidimensional integrals are involved in the likelihood function. Any increase in the number of observations over time increases that multi dimension. Also, the large number of parameters in the model makes the maximum likelihood estimates hard to compute. In the maximum likelihood estimation method, the inversion of square matrices with a dimension equal to the number of parameters are necessary. However, due to the large number of parameters, the inverse matrices can be inaccurate. Therefore, an approximate likelihood function is developed in the complex situation like this.

1.3 Review of Literature on Estimation of the Stochastic Factor Ordered-Probit Model

1.3.1 Granularity-based Estimation Method

The stochastic migration model is a dynamic (non-linear) risk factor model, that allows us to study the dynamics of credit rating matrices by introducing an unobservable common factor that represents a fundamental driving process. It is used for the prediction of future credit risk in a homogeneous pool of credits [see, e.g. Gupton et al. (1997); Gordy and Heitfield (2002); Gagliardini and Gouriéroux (2005); Feng et al. (2008)]. The ordered qualitative model with one-factor is a basic stochastic migration model and it is an extension of the Asymptotic Single Risk Factor (ASRF) model as well. The true log-likelihood function of the stochastic migration model involves a multi dimensional integral due to the presence of latent factors that need to be integrated out. Therefore, the computation of the true log-likelihood is complicated and time consuming [see, e.g. Schönbucher (2001, 2003); Frey and McNeil (2003) for a numerical implementation of the model].

Frye (2000) proposed an approximation of the true log-likelihood function for an application to the Asymptotic Single Risk Factor model (ASRF) under the assumption of infinitely large cohorts. It is called the Cross-Sectionally Asymptotic (CSA) log-likelihood function. The maximization of the (CSA) log-likelihood function helps to eliminate the multidimensional integral of the true log-likelihood and provides the CSA ML estimators. Although, the CSA approximation provides a simplified version of true log-likelihood and eases the procedure of estimation, the assumption of infinitely large cohort sizes is unrealistic and even inadequate in practice. Therefore, the assumption of infinitely large cohort sizes needed to be changed to large cohort sizes. A simplification based on the granularity principle which belongs to the class of analytical approximations of the true log-likelihood functions, provides an approximated log-likelihood for a large but not infinitely large cohort. Gagliardini and Gourieroux (2014b) illustrated the usefulness of this approach for a variety of problems related to risk analysis, statistical estimation, and derivatives pricing in finance and insurance. Gouriéroux and Jasiak (2012) provided the granularity adjustment for the CSA approximation based estimation of the true log-likelihood function of a default risk factor model under the assumption of large but not infinitely large cohort sizes.

The implementation of the granularity principle consists of three steps:

- In the first step, a risk factor model (RFM) is considered that connects the migration of individual credit ratings to systematic risk factors and unsystematic (i.e., idiosyncratic) risks. All transitions have a common source of uncertainty which is represented by systematic risk factors and cannot be diversified.

- Second, the RFM is extended to the asymptotic risk factor model (ARFM) under the assumption of an infinite portfolio of individual risks, which creates a simplification of the RFM model. The idiosyncratic risks will be completely diversified away in the ARFM model and only the systematic risk factors are left. With this simplification closed form formulas for the portfolio risk measures can be derived. In the setup of a portfolio of infinite size, that risk measure is called the cross-sectional asymptotic (CSA) risk measure .

- Third, the assumption of infinite portfolio size will be relaxed and an actual portfolio of a finite but large size will be considered. An asymptotic expansion around the ARFM will be performed to provide the closed-form approximation of the risk measure. This approximation consists of two parts: the approximation of the CSA and the adjustment term, called the granularity adjustment (GA). The diversified unsystematic risks in a virtual portfolio and their interaction with systematic risk factors in an actual portfolio are being taken care of by the GA part.

The non-linear risk factor model includes micro-dynamic and macro-dynamic components. The micro-dynamic component refers to the individual risk which is dependent on its own lagged values, conditional on the factor path. The macro-dynamic component refers to the common factor with serial persistence. The estimation of a stochastic migration model, which is a non-linear state space model is complicated. The data usually consists of a large panel of T time observations for n individuals. Therefore, the likelihood function involves a multi dimensional integral with respect to the unobserved factor and is hard to compute.

Gordy (2003), borrowed the terminology used in physics and photography and introduced the concept of granularity to economics and finance. A system is granular if composed of distinguishable pieces or grains or can be broken into small pieces or grains of similar size with no significant effect on the entire system. In the context of credit risk, a credit portfolio can be seen as a system and each individual loan defines a grain. Based on the granularity concept, a credit portfolio is granular, if each individual loan has the same exposure and no loans carry a systematic risk. The systematic risks along with unsystematic (idiosyncratic) risks are the components of the credit risk in portfolios. The systematic risks are not diversifiable and are the sources of uncertainty for all individual loans, as most borrowers are exposed to any unexpected changes in macroeconomic conditions and financial market conditions. Also, the systematic risks introduce dependencies between individual risks. The unsystematic risks (idiosyncratic) are all the remaining risks that can partly be diversified. The undiversified part of unsystematic risks can be assessed via the granularity adjustment methodology introduced by Gordy (2003).

Originally, the granularity theory was used to compute risks on large portfolios. Later on, this methodology became popular in different but related applications. For instance, it has been used for filtering out the unobservable values of systematic risk factors in a large panel of individual risk histories and to estimate the unknown parameters of a risk factor model. It has been used for pricing the derivatives as well. Moreover, the granularity adjustment (GA) methodology has been suggested by Basel II and Basel III regulation guidelines for computing the required capital reserve. Gagliardini and Gourieroux (2014b), have extended the granularity methodology further. A manageable framework provided by the granularity theory accommodates the effect of systematic risks, and idiosyncratic risks, the uncertainty about the model parameters, and the unobservability of states in risk analysis of large portfolios which makes this methodology important for financial applications.

Gagliardini and Gourieroux (2014b) used the granularity theory to derive asymptotically efficient estimators of parameters without computing the multi dimensional integrals of a complex log-likelihood. Gagliardini and Gourieroux (2014b) introduced computationally simple estimators for two kinds of parameters, i.e., the micro and macro-parameters in large dynamic panel models with common unobservable factors. This method is based on the assumption of infinitely large cohorts and is called the Cross-Sectionally Asymptotic (CSA) log-likelihood function. It is based on Frye (2000) where the Asymptotic Single Risk Factor model is considered and the CSA log-likelihood function is used to estimate the parameters of interest. Although a simplified estimation procedure can be applied to a CSA approximation to the true log-likelihood function, the assumption of infinitely large cohorts is unrealistic. Thus, it is more realistic and feasible to assume that the cross-sectional size is large, but finite, and use the granularity principle. This method provides closed-form approximations of the likelihood function for large cross-sectional dimension n followed by asymptotic expansions of the complicated likelihood function for panel models with a common latent factor and large cross-sectional dimension. There are two steps in the granularity method. The first step consist of virtual panels with infinite cross-sectional size and using the CSA approximation. The second step, assumes that the cross-section is large but finite and performs Taylor expansion of the likelihood function of order $1/n$ around the asymptotic model. This is the so called granularity adjustment (GA).

1.3.2 Maximum Composite Likelihood Estimation Method

Generally, there are many problems with statistical inference. When the number of data and parameters are huge (big data) or the data like credit rating changes have persistent discrete characteristics that make the modelling and estimating more challenging. In order to model such data with nonlinear features, we need a complex model that includes a likelihood function with multi dimensional integrals. This likelihood function is frequently very hard to compute in a parametric model. There exists methods, such as the Simulation Maximum Likelihood (SML), Beysian technique and Laplace approximation that help to compute multiple integrals and overcome this problem. Papers that used SML approach are for example, Hajivassiliou and Ruud (1994), Gourieroux et al. (1996), Lee (1997), and Fermanian and Salanie (2004) in the context of economics and Gagliardini and Gouriéroux (2005), Feng et al. (2008), Koopman et al. (2009), and Koopman et al. (2012) (Monte Carlo ML) in the context of finance. However, practitioners find this method complicated and possibly time-consuming as Feng et al. (2008) stated.

The maximum composite likelihood is a method which is not widely-known in practice but can be used when the computation of the full likelihood is infeasible due to underlying complex dependencies. There are two key incentives for using the maximum composite likelihood approaches. First, calculating the likelihoods for subsets of data is much easier than calculating the full-likelihood for the whole data set. Second, it neglects some dependence between the observations that makes the models complex. Generally, the composite likelihood can be defined by multiplying a collection of component likelihoods. The entire dependence can be ignored in the marginal composite likelihood methods [Chandler and Bate (2007)] and pairwise dependence can be captured in pairwise composite likelihood methods. [Cox (1975)]

The underlying idea behind the composite likelihood method is to replace the model by a misspecified likelihood. There exists different terminologies for composite likelihoods. For instance, Hens et al. (2005a), uses the term pseudo-likelihood or Stein et al. (2004) use approximate likelihood and the term quasi-likelihood is used by Hjort et al. (1994), Glasbey (2001), and Hjort and Varin (2008). Typically, there are two versions of composite likelihood, the conditional and the marginal likelihood. The product of the conditional densities of single observations given their neighbours is called the conditional pseudo-likelihood. Besag (1975) introduced the composite likelihood as a pseudo-likelihood in the context of models for spatially correlated data on a grid. Vecchia (1988) and Stein et al. (2004), use Besag's idea and develop a variant pseudo-likelihood method which involves chunks of observations in both the conditional and conditioned events. There exists further studies on Besag's proposal, like Hanfelt (2004a), Wang and Williamson (2005), and Fujii and Yanagimoto (2005) The second version of the composite likelihood, i.e, the composite marginal likelihood, is a pseudo-likelihood constructed under the independence assumptions. All serial dependence is ignored in this pseudo-likelihood. Chandler and Bate (2007) named this pseudo-likelihood, the independence likelihood. Under the independence assumptions, only inference on marginal parameters is possible in the independence likelihood. However, if dependent parameters are considered in the model, the pairwise likelihood introduced by Cox and Reid (2004) and Varin (2008) can be applied.

In the early 2000s, the composite likelihood approach became more popular in financial econometrics. Cox (1975) introduced the partial likelihood. Lindsay (1988), uses for the first time the term composite likelihood. Many papers in the statistics, biology and computer science have used the composite likelihood methods to estimate very complex systems. An overview of this methodology is given by Varin et al. (2011) in the context of statistical genetics. For instance, Pakel et al. (2017) used the composite likelihood for building models for multi dimensional portfolios. Gourieroux et al. (1984) analyze the asymptotic properties of the pseudo-maximum likelihood method. Fermanian and Salanie (2004) use non-parametric simulated maximum likelihood method to estimate parts of the full-likelihood of an autoregressive Tobit model. Lando and Skødeberg (2002) use a particular part of the likelihood function and estimate some parameters of their model by maximizing the partial-likelihood method.Vasdekis et al. (2012) analyze a dynamic factor structure in a probit model.

The composite likelihood method has a wide variety of applications in literature. For instance, studies which involve models that have dynamic latent variables, such as Probit models. An example is the AR-Probit model introduced by Varin and Vidoni (2006). That study is the first paper which applied the maximum composite likelihood to an AR-Probit model. Le Cessie and Van Houwelingen (1994) provided a pairwise pseudo-likelihood to estimate correlated binary data with an underlying process and Varin and Czado (2009), estimated the panel Probit model with autoregressive error structure by pairwise composite likelihood. Moreover, Ng et al. (2011) introduced some theoretical results of composite likelihoods in a more general model with a dynamic latent variable of stochastic volatility that follows an AR-Poisson process.

The composite likelihood can be written as a marginal or a conditional likelihood. These two versions of the composite likelihood are parts of the full-likelihood. Therefore, some theoretical properties of the maximum composite likelihood estimators will be close to the properties of the full-likelihood estimators. For instance, the maximum composite likelihood estimators are consistent, since, the marginal or conditional log-likelihood belongs in the full-likelihood. Then, Kullback and Leibler (1951) information inequality will hold for the maximum composite likelihood. The robustness of the composite likelihood inference has been investigated by many authors. Most of these studies refer to the maximum composite likelihood inference as robust, since the assumptions of maximum composite likelihood inference are based on lower dimensional conditional or marginal densities, instead of detailed specifications of the full joint distribution. There are some studies like, Lele and Taper (2002) and Wang and Williamson (2005) that examine the robustness of the maximum composite likelihood inference in more details.

Concerning the efficiency, the asymptotic variance of a simplified likelihood estimator or "miss-specified likelihood" estimator, is higher than that of the full-likelihood estimator. The inverse of information matrix which is the asymptotic variance of the full-likelihood does not hold for the maximum composite likelihood. Instead, the asymptotic variance of the composite likelihood is the so-called sandwich form or the Godambe Information [Godambe (1960)]. Establishing the general efficiency result for the composite likelihoods is hard. Mardia et al. (2007) shows that composite conditional likelihoods (marginal and conditional) are fully efficient for the general multivariate normal distribution in applications to highly concentrated i.i.d data. Lindsay et al. (2011) develop a methodology to increase the efficiency of the maximum composite likelihood estimator. Consequently, a weighting scheme for the maximum composite likelihood has been proposed by Harden (2013). However, these studies show low efficiency of the maximum composite likelihood estimators. The efficiency loss of a maximum composite likelihood estimator in the ARFIMA models is evaluated in Davis and Yau (2011). They show that the maximum composite likelihood estimator loos efficiency if the order of integration is high in a long-memory process. Varin and Vidoni (2006) and Joe and Lee (2009), analyzed the efficiency loss in the context of time series and discover sign of improvement in efficiency if only nearly next to pairs are included in the maximum composite likelihood. Concerning the identification of parameters in maximum composite likelihood, the identification of model parameters may not be achieved, depending on which components are included in the composite likelihoods. Especially in more complex model, one can not identify all the parameters of the model.

The alternatives estimation techniques such as the simulated maximum likelihood (SML), the Beysian estimation and Laplace approximation are often complicated and possibly time-consuming. Thus, a two stage estimation where, in the first stage, the model is estimated without the dynamic and next in stage two the dynamic is estimated has been offered in literature. A comparison of the performance of the maximum composite likelihood and the simulation maximum likelihood methods in multivariate ordered-response situations has been done by Bhat et al. (2010).

Their study shows that both maximum composite likelihood and SML are efficient but in terms of computational time, the maximum composite likelihood is 40 times faster than the SML. In the case of a Beysian method, there are studies which show the computational complexity as well. For instance, Chauvet and Potter (2005), McNeil and Wendin (2007), and Stefanescu et al. (2009), consider a Probit model with a latent factor and use Gibbs sampling to estimate it. A sophisticated approach is introduced by Müller and Czado (2005) since Gibbs sampling exhibits bad convergence. However, Varin and Vidoni (2006) questioned this proposed approach due to increasing computational complexity. Moreover, a Taylor approximation of the likelihood has been proposed by Gagliardini and Gourieroux (2014a).

1.4 Description of Credit Ratings Dynamics

1.4.1 Individual Credit Rating Histories

Individual firm (i) has a credit rating history which is denoted by $(y_{i,t},\ t = 1, ..., T)$, $i = 1, ..., n$ and corresponds to the qualitative process with state space $(1, ..., K)$ at any point of time t. These states indicate all possible rating categories such as $A^+, A, ..., F$ of firm $i = 1, ..., n$ and characterize different risk categories. State (A^+) refers to the low risk category and (F) refers to state default, which is an absorbing state and the highest risk category. The qualitative ratings $y_{i,t}$ are determined from an unobserved continuously valued score $y^*_{i,t}$. The qualitative measure of risk $y_{i,t}$ which is called the rating is compatible with a scoring system and by discretizing the score grade, the rating can be defined. The history of credit ratings of a given borrower can be used for:

- Approximating the score function y^*_{it}. The sequence of previous ratings is used in scoring practice to approximate the credit scores of a given borrower.

- Predicting the risk on credit portfolio from the information of past individual credit ratings.

In both cases the serial correlation in observed ratings matters. In order to be able to analyse the effect of lagged ratings on current ratings, the serial correlation should be considered. Moreover, the cross-sectional correlation among the observed ratings is important for analysing the migration correlation, which shows the relation between joint upgrades and downgrades of different firms as well. However the latter is not the attention of this chapter. This chapter explains the dynamics of rating transitions and captures the serial and cross-sectional correlation with an unobservable time varying stochastic factor that represents a common risk factor that affects all individual ratings at each time $t = 1, ..., T$.

1.4.2 Observed Transition Matrices

Credit ratings of an individual firm (i) change over time. Various approaches can be distinguished with the respect to the assumption on P_t and y^*_{it}. A study of rating changes of a firm can help forecast the future risk. A history of ratings is the basic element in defining the matrices of credit rating transitions P_t. The standard approach to define the sequence of credit rating transitions matrices is to aggregate the individual firm credit rating histories into migration frequencies. The assumption for the basic approach to define the sequence of transition matrices is as follows:

Assumption 1.1- Independent, identically distributed (i.i.d.) process of credit rating histories across the firms $(y_{it},\ t = 1, ..., T)$, $i = 1, ..., n$ and each individual credit rating history follows a heterogeneous Markov process.
The heterogeneous Markov process means that all the information on the rating history of a given firm can be explained by the most recent individual rating and that the migration frequencies can be time varying. The transition frequencies are the same for different individuals, which follows from the cross-sectionally homogeneity assumption on the population of firms. Each element of these matrices refers to frequencies of transitions from one rating category to another in a given period of time in a given group of bond issuers. The entries

of each matrix are positive and each row sums up to one, since they represent the probabilities of migration. The dimensions of each matrix depends on the number of states (k). The matrices are updated and reported in each year by the rating agencies. The analysis of such matrices can be performed either under the assumption that $P_t,\ t = 1, ..., T$ are time independent (time homogeneity assumption) or time varying (heterogeneity assumption). In the later case, transition matrices are assumed dependent on time, and can be modelled by either the deterministic or stochastic models.

Based on assumption 1.1, the distribution of each rating history is characterized by a sequence of migration matrices $P_t,\ t = 1, ..., T$ and the sequence of migration matrices describes the movement of each firm ratings from one rating category to another. The dimension of the transition matrices is $(K \times K)$ and each entry displays the probability of changing in firm's rating from (l) to (k) between time $(t-1)$ and (t).

$$p_{k,l,t} = P[y_{i,t} = k | y_{i,t-1} = l],\ \forall\ k, l, t. \tag{2.1}$$

Deterministic Transition Matrices: The assumptions 1.1 is the main assumption in deterministic transition model. Since the default is absorbing states and once one firm's rating become default, it will stay there, and it can not exit the default state, each matrix, is reduced to K columns and $(K-1)$ rows. The probability of transition from rating l to default K is as follows:

$$p_{k,K,t} = \begin{cases} 1 & \text{if k=K} \\ 0 & \text{otherwise} \end{cases}$$

Therefore, there are $(K-1)^2$ independent elements in each transition matrix that depend on the number of parameters in the model that should be estimated. Based on assumption 1.1, it can be concluded that the transition probabilities for different individuals are equal since population of firms is homogeneous. However, it is not always the case that the homogeneity of firms is assumed. For instance, Gordy (2003) and Gagliardini and Gouriéroux (2005) questioned the population homogeneity assumption in their studies. The following three criteria: industrial sector, country of origin and credit quality rating that are used to define a homogeneous population by Basel Committee on Supervision.

Stochastic Transition Matrices: In stochastic transition model, the migration matrices P_t are stochastic. The sequence of transition matrices is influenced by factors. These factors can be divided into observable and unobservable factors. Two important correlations, the cross-sectional correlation between risk and serial dependence in risk will be captured when the factors are unobservable. In this model the assumption of deterministic model has to be modified as follows:

Assumption 2.1- The vector of individual risk histories $\{y_{i,t}\}$, for $i = 1, ..., N,\ t = 1, ..., T$ and follow a heterogeneous Markov process with transition matrices $\{P_t\}$.

Assumption 2.2- The sequence $\{P_t\}$ is stochastic.

There are two major advantages for assuming stochastic transition matrices. Since the stochastic dynamics involves a small number of underlying factors, the ratings model is flexible. The heterogeneous Markov process also can be used to predict the future risk if the dynamics of migration matrices are well-defined. Also, in order to capture the joint upgrades and downgrades of firms, stochastic models can be used to measure migration correlation. Generally, the stochastic models are more flexible and useful for modelling default.

1.5 The Ordered Qualitative Variable Model

We use an ordered qualitative dependent variables framework to model the credit rating transition matrix, due to the fact that there is a natural ordering from low to high credit quality, in credit ratings. The ordered-Probit model belongs to the family of ordered qualitative variable models. The ordered-Probit model is obtained when the error terms follow the standard Gaussian distribution. In this model, there is a latent continuous quantitative score $y^*_{i,t}$ that determines the individual qualitative ratings. This unobservable score y^*_{it} is discretized in order to obtain the individual qualitative ratings. Moreover, there is another latent variable, called the factor that drives the parameters of the latent score and make them stochastic.

In the stochastic transition matrices setup, we introduce an unobserved common factor that drives the migration probabilities. Thus the transition probabilities depend on a factor f_t and fixed parameters $\tilde{\theta}$:

$$p_{k,l,t} = p_{kl}(f_t; \tilde{\theta}),$$

$$\theta_t = f(f_t, \tilde{\theta}).$$

In the next section, the stochastic factor ordered-Probit model known as a stochastic migration model is explained in detail.

1.5.1 Stochastic Migration Model

The effect of current ratings of borrowers on the value of their debts in a credit portfolio of a financial institution has to be taken into account when the risk on a credit portfolio is assessed. An accurate analysis of risk associated with possible rating downgrades and upgrades is a requirement of the Basel 2 accord [Basel Committee on Banking Supervision (2001, 2003)]. Thus, it is required to perform a dynamic analysis of qualitative rating histories. In order to model the dynamics of a large number of qualitative individual histories $(y_{i,t})$ over the finite state $(1, ..., K)$, we use the stochastic transition model.

The object of our analysis is a sequence of observed square matrices $\hat{P}_t$ with positive elements, which are the frequencies of transitions between states. Each element of these matrices refers to frequencies of transitions from one rating category to another in a given period of time in a given group of bond issuers. The elements of each matrix are positive and each row sums up to one, since they represent the probabilities of migration. The dimensions of each matrix depends on the number of states k. The matrices are updated and reported each year by the rating agencies. The analysis of such matrices can be performed either under the assumption that $\hat{P}_t,\ t = 1, ..., T$ are time independent (time homogeneity assumption) or time varying (heterogeneity assumption). We consider stochastic time varying matrices in our analysis. Let us explain the model in more details.

Individual firm (i) has a credit rating history which is denoted by $(y_{i,t},\ t = 1, ..., T),\ i = 1, ..., n$ and corresponds to the qualitative process with state space $(1, ..., K)$ at any point of time t. These states indicate all possible rating categories such as $A^+, A, ..., F$ of firm $i = 1, ..., n$ and characterize different risk categories. State (A^+) refers to the low risk category and (F) refers to state default, which is an absorbing state and the highest risk category. The qualitative ratings $y_{i,t}$ depend on an unobserved continuously valued score $y^*_{i,t}$. We introduce two main assumptions:

Assumption 3.1- The individual risk histories $\{y_{i,t}\}$, $i = 1, ..., N$, $t = 1, ..., T$, are i.i.d. across the firms. Each individual conditioning history follows a heterogeneous Markov process with transition matrices $\{P_t\}$.

Assumption 3.2- The sequence $\{P_t\}$ is stochastic.

The heterogeneous Markov process means that all the information on the rating history of a given firm can be explained by the most recent individual rating and that the migration probabilities are time varying. The transition probabilities are the same for different individuals, which follows from the cross-sectional homogeneity assumption on the population of firms.[2]

Based on assumption 3.1 and 3.2 the rating history is characterized by a sequence of migration matrices P_t, $t = 1, ..., T$ and the sequence of migration matrices describes the movement of each firm ratings from one rating category to another. The dimension of the transition matrices is $(K \times K)$ and each entry displays the probability of changing in firm's rating from (l) to (k) between time $(t-1)$ and (t).

$$p_{k,l,t} = P[y_{i,t} = k | y_{i,t-1} = l, f_t], \ \forall \ k, l, t. \tag{1.5.1.1}$$

By discretizing the score grade, the qualitative rating of corporates can be defined, as follows. The quantitative values of score y^*_{it} can be between $c_0 = -\infty < c_1 \leq ... \leq c_K = \infty$. Based on the score value, the ratings are defined by:

$$y_{i,t} = k, \quad \text{iff} \quad c_{k-1} \leq y^*_{i,t} < c_k, \quad k = 1, ..., K. \tag{1.5.1.2}$$

The rating classes are numbered in order of increasing credit quality, with alternative $k = 1$ corresponding to the highest risk category or default (F) and $k = K$ corresponding to the

[2]See Gagliardini and Gourieroux (2014b), Definition 1.1, Page 6.

lowest risk category ("A+"). The movements of firms from one rating category to another depends on a common unobserved factor f_t:

$$y^*_{i,t} = \delta_l + \beta_l f_t + \sigma_l u_{i,t}, \quad \text{if} \ \ y_{i,t-1} = l, \ l = 1, ..., K, \tag{1.5.1.3}$$

where $u_{i,t} \sim IIN(0, 1)$ is independent of f_t. By conditioning with respect to the last rating we obtain a set of K homogeneous sub-populations at date t. It means that firms are grouped into rating categories $k = 1, ..., K$ according to their previous rating. It is important to specify the conditional distribution $\psi(f_t|f_{t-1}; \rho)$. The unobserved common factor satisfies a Gaussian Autoregressive (AR(1)) model:

$$f_t = \mu + \rho f_{t-1} + \eta_t, \qquad \eta_t \in IIN(0, \sigma^2_\eta), \tag{1.5.1.4}$$

and the errors u_{it}, η_t are identically independent. We assume of $\rho \in (-1, 1)$, to ensure a stationary behaviour of the factor. Thus, our dynamic model is defined by a state and measurement equations as follows:

State equation: It is defined by the conditional density of the factor: $\psi(f_t|f_{t-1}; \rho), \ \forall t$

Measurement equations: These equations are defined by the transition probabilities:

$$\begin{aligned} p_{lk,t}(f_t; \theta) =& P[y_{i,t} = k|y_{i,t-1} = l, f_t; \theta] \\ =& P[c_{k-1} \leq y^*_{it} < c_k|y_{it-1} = l], \end{aligned} \tag{1.5.1.5}$$

for $l, k = 1, ..., K$, where θ is a vector of micro-parameters defined below.

The transition density function of $y_{i,t}$, given $y_{i,t-1}$ and f_t is characterized by the $(K \times K)$ migration matrix, since the individual observations are qualitative. Under equations (1.5.1.2), (1.5.1..3) and (1.5.1.5) the dynamic model is a Probit with the transition probabilities given by:

$$p_{lk,t}(f_t; \theta) = \Phi(\frac{c_k - \delta_l - \beta_l f_t}{\sigma_l}) - \Phi(\frac{c_{k-1} - \delta_l - \beta_l f_t}{\sigma_l}), \ l, k = 1, ..., K, \ t = 1, ..., T, \tag{1.5.1.6}$$

where Φ is the standard normal cumulative distribution function (cdf) in this framework [3]. There are two kinds of parameters: the micro-parameters and macro-parameters:

[3] Each row of the transition matrix, conditional on (f_T) contains an ordered-Probit model with a common factor (f_t). When factoris serially correlated, as in (1.5.1.4), the transition matrices are serially correlated.

Micro-parameters- The parameter vector θ includes the thresholds c_k, $k = 1, ..., K-1$ and parameters of latent score function, $(\delta_l, \beta_l, \sigma_l),\ l = 1, ..., K$.

Macro-parameter- Parameter ρ characterizes the factor dynamics.

The vector of micro-parameters θ characterize the micro-dynamics of ratings and macro-parameter ρ characterizes the factor dynamic. The micro-dynamics are defined by the measurement equation and the macro-dynamics are defined by the state equation and are both conditional on a given factor path. Moreover, the individual histories of the credit ratings of all firms are influenced by a common factor f_t and this common factor creates cross-sectional dependence between individual histories. Before applying the estimation method, we need to consider the parameter identification.

1.5.2 Identification

Identification constraints need to be imposed on the micro- and macro-parameters of a stochastic factor ordered-Probit model. [4] The identification constraints create the links between the rows of the transition matrix. There are two reasons for that lack of identification. First, in the process of moving from the quantitative scoring variable $(y^*_{i,t})$ to the qualitative rating $(y_{i,t})$, some information is lost. Second, the factor is unobserved and defined by an invertible linear transformation. Thus, the identifying constraints can be as follows:

First, due to an invertible affine transformation of the factor, the identifying constrains can be imposed on factor dynamics:

$$E(f_t) = 0, \text{ and } V(f_t) = 1.$$

Second, since, the quantitative score is partially observable, therefore, various combinations of affine transformations of the quantitative score and of the thresholds can provide the same migration probabilities. So, it is sufficient to impose the standard identification

[4] A vector of model parameters is not identified, if we can find two different parameter vectors with the same distribution for the observable variables, and the joint history of the ratings of the n firms. [see, e.g., Gourieroux and Monfort (1995) and Gagliardini and Gouriéroux (2010)]

restrictions for an ordered-Probit model on one row only, for instance:

$$\delta_l = 0 \text{ and } \sigma_l = 1, \quad \text{or} \quad c_k = 0 \text{ and } \sigma_l = 1.$$

The identification restrictions on the parameters and factors depend on the estimation method [see, Gagliardini and Gouriéroux (2005)]. The identification constraints for each estimation methods are discussed in more detail in the next two chapters.

1.6 Statistical Inference

1.6.1 Estimation of the Stochastic Factor Model

In general, when the transition matrices are deterministic, the maximum likelihood is the best methodology for estimating the parameters of interest in the model. However, in the case of stochastic transition matrices, the maximum likelihood approach is hard to implement, due to computational complexity. The stochastic migration model is considered as a special case of multi-factor model for panel data. In such a framework, the likelihood function involves multidimensional integrals of multi dimension that makes the maximum likelihood or GMM approaches intractable and hard to compute.

Let us explain how the likelihood function is defined in the stochastic migration model. In the stochastic model, the migration matrices are functions of the parameters of score function as well as the common factor value (f_t). The vector θ includes all parameters in the state space model, that are: $\delta_l, \beta_l, \sigma_l,\ l = 1, ..., K$, corresponding to the quantitative score, and parameters c_k, $k = 2, ..., K$ defining the state discretization. Since the conditional migration matrices are functions of the parameters θ as well as of the common factor values (f_t), the conditional likelihood function is:

$$L(\underline{y_T}|F, y_1; \theta) = \prod_{t=1}^{T} \prod_{k=1}^{K} \prod_{l=1}^{K} (P_{kl}(f_t; \theta))^{n_{klt}}, \tag{1.6.1.1}$$

where n_{klt} denotes the number of firms which migrate from l to k between $t-1$ and t. The vector of the individual histories $\underline{y_T} = y_{it}$ for $i = 1, ..., n$ and $t = 1, ..., T,\ F = (f_t)$,

$t = 2, ..., T$ and $y_1 = (y_{11}, ..., y_{n1})'$. Since both rating and factor histories are not observed, the distribution of factor $(f_1, ..., f_t)$ values have to be integrated. Therefore, the true likelihood function is given by:

$$L(\underline{y_T}|F, y_1; \theta) = \int ... \int \left(\prod_{t=1}^{T} \prod_{k=1}^{K} \prod_{l=1}^{K} (P_{kl}(f_t; \theta))^{n_{klt}} \psi(f_1, ..., f_T; \rho) df_1 ... df_T \right).$$

After taking the logarithm, it becomes:

$$\log\ L(\underline{y_T}|F, y_1; \theta) = l(\underline{y_T}|y_1; \theta, \rho),$$

$$l(\underline{y_T}|y_1; \theta, \rho) = \log\ \int ... \int \left([\prod_{t=1}^{T} \psi(f_t|f_{t-1}; \rho)] \prod_{t=1}^{T} \prod_{i=1}^{n} h(y_{i,t}|y_{i,t-1}, f_t; \theta) \prod_{t=1}^{T} df_t \right), \quad (1.6.1.2)$$

where ψ refers to joint distribution of factors. The log-likelihood function involves a multivariate integral with dimension equal to the time T, which is usually large and hard to compute in practice. The dimension of this integral increases with T. Usually, practitioners use variational approximation, Simulated Maximum Likelihood (SML), quasi-likelihood, Indirect Inference, and Laplace-type approximations. Feng et al. (2008) use the Simulated Maximum Likelihood (SML) and replace the multivariate integrals by an approximation which is computed by simulations. [see, Gourieroux et al. (1996)] There exists several algorithms in statistical software packages that are able to numerically optimize this complicated likelihood function, see e.g. Frey and McNeil (2003) and Hamerle and Rösch (2005, 2006). However, this procedure is quite complicated and time consuming.

1.7 Conclusion

In this chapter, we presented the literature review on the dynamics, estimation and modelling of credit rating transition matrices. We explained the general framework of the stochastic migration model. We also discussed the statistical inference of the latent factor ordered-Probit model. Due to computational complexity of the full information maximum likelihood in the latent factor-Probit model, this thesis three maximum likelihood estimation methods of the latent factor ordered-Probit model. The first two methods rely on the Asymptotic Risk Factor (ARF) model under the assumption of an extremely large portfolio. These approaches belong to the class of analytical approximations that are derived from the so-called granularity principle or granularity theory. The third methodology is the maximum composite likelihood which is borrowed from the statistical literature and offers faster and more robust estimation results. This method disregards some of the complex dependencies between observations in the full joint model. For instance, the joint densities can be replaced by a product of pairwise joint densities or a product of conditional densities. Therefore, in the composite likelihood method, the objective function is derived by multiplying a collection of component likelihoods. These simplified likelihood methods offer fast and reliable estimation that provides the output quickly. The first two methods, the two-step and the joint optimization are discussed in Chapter 2 and the third method, the maximum composite likelihood is discussed in Chapter 3.

Chapter 2

Granularity-based Estimation Method of the Stochastic Migration Model

2.1 Introduction

Financial institutions hold large portfolios of individual risks, such as stocks and corporate bonds, as well as individual contracts, such as corporate loans, household mortgages, and life insurance contracts. The size of these portfolios, the nonlinearities of risk models, and the dependencies between individual risks make the process of risk analysis very complex. The nonlinearities in the portfolios of individual risks are induced by the risks associated with rating migrations, default, and prepayment of credit portfolios. The individual securities and contracts become dependent due to the systematic risk factors that affect the payoffs of individual assets. In the literature, simulation-based methods for computing the risk measures, and estimating the risk models have been proposed. However, they can be very time consuming and can be inefficient in some cases like, the computation of capital reserves or investigating the effects of model parameters and risk factors on portfolio risk. Therefore, analytical approximations of the risk measures become of interest. One method of obtaining that approximated value relies on the Asymptotic Risk Factor (ARF) model, which provides the value of a risk measure under the assumption of infinity large portfolios. Another approximation approach exists, which relaxes the assumption of infinity large portfolios and provides the approximated value of a risk measure for large but not infinity large portfolios. It is based on the so-called granularity principle. The Basel Committee on Banking Supervision (BCBS) suggests banks to use this methodology for their credit risk analysis and their computation of capital reserves in an early draft of Basel II. [see, e.g. Basel Committee on Banking Supervision (2001); Wilde (2001); Martin and Wilde (2002); Gordy (2003, 2004); Gagliardini et al. (2012); Gordy and Lütkebohmert (2007); Gagliardini and Gouriéroux (2013)].

Gagliardini and Gourieroux (2014b) illustrated the usefulness of this approach for a variety of problems related to risk analysis, statistical estimation, and derivatives pricing in finance and insurance. Gouriéroux and Jasiak (2012) provided the granularity adjustment for the CSA approximation based estimation of the true log-likelihood function of a default risk factor model under the assumption of large but not infinitely large cohort sizes.

In this chapter, we follow Gagliardini and Gourieroux (2014b), and explore three estimation methods for the stochastic migration model. In the first approach, we rewrite the stochastic migration model as an approximate linear state space model, as described in Gagliardini and Gouriéroux (2005), and estimate it by means of the standard Kalman filter. This approach provides consistent and fully efficient estimators of the parameters of the linear model and also an approximated factor path when the cross-sectional dimension n, i.e. the number of firms in the sample, is large. The path of the factor is obtained exogenously from the principal component analysis in this approach. If we do not have a sufficiently large number of firms in a sample, the method loses its optimality. However, this method can be used to determine the number of factors. Also, the parameter estimates can be used as initial values in a numerical algorithm to maximize the approximated log-likelihood of the model of interest under the granularity principle, such as the two-step efficient estimation approach, which is the method discussed next.

In the second approach the stochastic migration model is estimated by the two-step estimation method which is the simplified likelihood based on the granularity theory, which provides asymptotically consistent estimators. The parameter set is divided into micro-parameters that determine the credit ratings and macro-parameters that define the factor dynamics. The estimates of micro- and macro-parameters are derived in two steps. In the first step, the factor values are considered as nuisance parameters and the estimator works approximately as a fixed effects estimator. In the second step, the unobserved factor values are replaced by cross-sectional factor approximations and the estimator of the micro-parameters is obtained. Next, the unobserved factor values are replaced by cross-sectional factor approximations and the estimator of the macro parameters can be obtained either by maximizing the likelihood of the macro-dynamics or by applying the maximum likelihood estimation of a autoregressive AR(1) model of factor dynamics.

The third method is the joint optimization approach where the micro-parameters and the factors are all considered as the parameters in the model. The estimates of the micro- and macro-parameters can be obtained by considering the solution of the joint optimization problem. We use the estimates of the two-step estimation approach as initial values in the numerical algorithm to maximize the log-likelihood.

This chapter empirically compares these three statistical estimation methodologies in application to the stochastic migration model, which links the transition probabilities to an unobserved dynamic risk factor. We use the same data set in each approach to examine each estimation method. In the two-step approach, we use the estimates of the approximate linear state-space model as initial values in the numerical algorithm to maximize the log-likelihood function.

The chapter is organized as follows. Section 2.2, introduces the approximate linear state-space model, the two-step efficient estimators and the joint optimization. Section 2.3 explains the data set. In Section 2.4, the estimation results of these three methods are discussed. Section 2.5 presents the stress test based on applying one-time shocks to the factor. The conclusion is given in section 2.6.

2.2 The Likelihoods-Granularity-Based Estimation Approach for the Stochastic Migration Model

In this section we discuss the likelihood-based estimation methods for the latent factor ordered-Probit model discussed in Section 1.5.1 These methods are based on analytical approximations of the log-likelihood function that are derived from the so-called granularity principle.

2.2.1 Approximate Linear State Space Model

The stochastic migration model can be represented as an approximate linear state space model. In this approach, for each row of the transition matrix, a canonical factor is computed. An estimation approach which relies on this linearized qualitative model is straight forward. Let us recall the equations (1.5.1.5) and (1.5.1.6).

$$\begin{aligned} p_{lk}(f_t;\theta) &= P[y_{i,t} = k | y_{i,t-1} = l, f_t; \theta], \\ &= \Phi(\frac{c_k - \delta_l - \beta_l f_t}{\sigma_l}) - \Phi(\frac{c_{k-1} - \delta_l - \beta_l f_t}{\sigma_l}), \quad k = 1, ..., K, \quad t = 1, ..., T. \end{aligned}$$

We deduce the cumulative transition probabilities p^*_{lk} for each row as follows:

$$\begin{aligned} p^*_{lk}(f_t;\theta) &= P[y_{i,t} \leq k | y_{i,t-1} = l, f_t; \theta], \\ &= \sum_{h=1}^{k} p_{lh}(f_t;\theta), \qquad (2.2.1.1) \\ &= \Phi(\frac{c_k - \delta_l - \beta_l f_t}{\sigma_l}), \end{aligned}$$

for $l = 1, ..., K$ and $k = 1, ..., K-1$. If we apply the quantile function of the standard normal distribution to both sides of equation (2.2.1.1) we get:

$$\Phi^{-1}[p^*_{lk}(f_t;\theta)] = \frac{c_k - \delta_l - \beta_l}{\sigma_l} = \frac{c_k - \delta_l}{\sigma_l} - \frac{1}{\sigma_l}\beta_l f_t. \qquad (2.2.1.2)$$

These nonlinear transformation of the cumulative transition probabilities for $l = 1, ..., K$ and $k = 1, ..., K-1$, play the role of the canonical factors which are linear with regard to f_t, and are denoted as:

$$a_t = vec[a_{l,k,t}], \tag{2.2.1.3}$$

where

$$a_{lk,t} = \frac{c_k - \delta_1 - \beta_l f_t}{\sigma_l} = \frac{c_k - \delta_l}{\sigma_l} - \frac{1}{\sigma_l}\beta_l f_t. \tag{2.2.1.4}$$

The probability $p^*_{lk,t}$ is well approximated by its cross-sectional sample counterpart when the cross-sectional dimension n is large $(n \to \infty)$. Furthermore, the estimated canonical factors $\hat{a}_{l,k,t}$ are such that

$$\hat{a}_{lk,t} = \Phi^{-1}\Big(\sum_{h=1}^{k} \hat{p}_{lh,t}\Big), \tag{2.2.1.5}$$

Therefore, for a large cross-sectional dimension n, we get the estimates of canonical factors $\hat{a}_t$, $t = 1, ..., T$, which are asymptotically normally distributed with their estimated asymptotic variance-covariance matrix $\hat{\Sigma}_{n,t}$,

$$\hat{a}_t \overset{d}{\sim} N(a_t, \hat{\Sigma}_{n,t}).$$

The derivation of the asymptotic variance-covariance matrix $\hat{\Sigma}_{n,t}$ is given in [Gagliardini and Gourieroux (2014b)]. The state equation and the measurement equation can be written as follows:

State equation:

$$f_t = \mu + \rho f_{t-1} + \eta_t, \quad \eta_t \sim IID(0,1),\ t = 1, ..., T.$$

Measurement equation:

$$\hat{a}_{lk,t} = \frac{c_k - \delta_l}{\sigma_l} - \frac{\beta}{\sigma_l} f_t + u_{lk,t}, \quad vec(u_{lk,t}) \sim IIN(0, \hat{\Sigma}_{n,t}),$$

for $l = 1, ..., K$ and $k = 1, ..., K$.

In this approach, the nonlinear measurement equations (1.5.1.5) are approximated by the linear measurement equations for a sufficiently large cross-sectional dimension n and can be analyzed by the standard linear Kalman filter. The estimates of parameters c, β, δ, and σ can be obtained by the standard Kalman filter under the following identification restrictions:

First, due to an invertible affine transformation of the factor, the identifying constraints can be imposed on the factor dynamics:

$$E(f_t) = 0, \text{ and } V(f_t) = 1.$$

Second, since, the quantitative score is partially observable, we need to impose the standard identification restrictions for an ordered-Probit model for one row only. The identification restrictions for micro-parameters in this model concern the parameters of rating class "B":

$$\delta_4 = 0 \text{ and } \sigma_4 = 1.$$

After imposing above restrictions on the factor dynamics, the micro-parameters, and given $E(f_t) = 0$, all remaining parameters can be obtained from:

$$\frac{c_k - \alpha_l}{\sigma_l} = \frac{1}{T} \sum_{t=1}^{T} \hat{a}_{lk,t},$$

$$\frac{\beta_l}{\sigma_l} = \text{Standard error of } \hat{a}_{lk,t},$$

for $l = 1, ..., K$ and $k = 1, ..., K - 1$.

If the cross-sectional dimension n tends to infinity, the errors in the measurement equations tends to zero. Therefore, we get an approximated linear state space model in which the macro-component is determined from the transition equation and the micro-components from the measurement equation. The standard linear Kalman filter can be applied to estimate this approximated model under the identification restrictions. In this approach, the rate of convergence for the micro-parameters is $\frac{1}{\sqrt{nT}}$ and for the macro-parameter is $\frac{1}{\sqrt{T}}$. This approach provides estimates of the micro-parameters, c, δ, β, σ and also allow us to filter the latent factor values. The factors can be obtained by the spectral decomposition of the $(T \times T)$ matrix YY', where the row t of matrix Y is given by $(\hat{a}_{lk,t} - \bar{a}_{lk})$, with, $(\bar{a}_{lk} = \frac{1}{T} \sum_{t=1}^{T} a_{lk,t})$. In order to illustrate the estimation procedure, let us consider an example of the model with three rating categories $k = 3$ and discuss the estimation steps.

Illustration: Qualitative Stochastic Transition Model

An illustration of the approximate linear state-space model outlined in the Section (2.2.1) is explained in this section. We consider a stochastic migration model with three $K = 3$

rating categories $k = 1, ..., 3$. The individual rating $y_{i,t}$ of firm $i = 1, ..., n$ can be in either qualitative rating class "1" , "2" , or "3" at time $t = 1, ..., T$. The class "1" corresponds to the high risk category of default and "3" refers to the lowest risk category. There is a pool of homogeneous corporates. Each of them has a credit rating at time $t-1$ and by the time t, either one of three transitions can occur:

- Firm i stays in the same rating class
- Firm i is upgraded to a higher rating class
- Firm i is downgraded to a lower rating class

Table 2.1 provides the counts of firms in each class, which is the main summary that can be used to calculate the transition frequencies for each rating category as follows:

Table 2.1: Number of Firms in Each Rating Class

(t-1)	(t)			
	3	2	1	Issuers
3	$n_{33,t}$	$n_{32,t}$	$n_{31,t}$	$n_{3,t-1}$
2	$n_{23,t}$	$n_{22,t}$	$n_{21,t}$	$n_{2,t-1}$
1	$n_{13,t}$	$n_{12,t}$	$n_{11,t}$	$n_{1,t-1}$
Total	$n_{3,t}$	$n_{2,t}$	$n_{1,t}$	

Thus, the transition matrix can be calculated as follows:

Table 2.2: Transition Matrix for Year t. $\hat{p}_{lk,t}$ is observed transition frequency from category l to k at time t.

	3	2	1	Sum
3	$(n_{33,t})/(n_{3,t-1})=p_{33,t}$	$(n_{32,t})/(n_{3,t-1})=p_{32,t}$	$(n_{31,t})/(n_{3,t-1})=p_{31,t}$	1
2	$(n_{23,t})/(n_{2,t-1})=p_{23,t}$	$(n_{22,t})/(n_{2,t-1})=p_{22,t}$	$(n_{21,t})/(n_{2,t-1})=p_{21,t}$	1
1	$(n_{13,t})/(n_{1,t-1})=p_{13,t}$	$(n_{12,t})/(n_{1,t-1})=p_{12,t}$	$(n_{11,t})/(n_{1,t-1})=p_{11,t}$	1

The transition matrices $\hat{P}_t = [p_{lk}]_t$ are observed at each time t. Their dimensions are $(K \times K)$ and each P_t has K^2 elements, $k = 1, ..., k$. The variation of theses matrices over time is due to the latent factor f_t. The firm i gets the rating in class k, where $k = 1, 2, 3$ if the latent score y^*_{it} lies between threshold $c's$:

$$y_{i,t} = k \text{ if } c_{k-1} < y^*_{i,t} < c_k.$$

Since there are three rating categories, we have two thresholds, denoted by c_1 and c_2. Thus we have:

$$\begin{aligned} y_{i,t} &= 1 \;\text{ if } y^*_{i,t} \leq c_1, \\ y_{i,t} &= 2 \;\text{ if } c_1 < y^*_{i,t} \leq c_2, \\ y_{i,t} &= 3 \;\text{ if } y^*_{i,t} > c_2, \end{aligned}$$

by assuming that the latent scores are linear functions of factor f_t, the linear factor model can be written as follows:

$$\begin{aligned} y_{i,t-1} &= 1 \rightarrow y^*_{i,t} = \delta_1 + \beta_1 f_t + \sigma_1 u_{i,t}, \\ y_{i,t-1} &= 2 \rightarrow y^*_{i,t} = \delta_2 + \beta_2 f_t + \sigma_2 u_{i,t}, \\ y_{i,t-1} &= 3 \rightarrow y^*_{i,t} = \delta_3 + \beta_3 f_t + \sigma_3 u_{i,t}, \end{aligned}$$

where the factor f_t satisfy a Gaussian Autoregressive (AR(1)) process. The error terms $u_{i,t}$ are i.i.d normal with mean zero and variance one. There are 9 unknown parameters $(\delta_1, \delta_2, \delta_3, \beta_1, \beta_2, \beta_3, \sigma_1, \sigma_2,$ and $\sigma_3)$ and two thresholds (c_1, c_2) and T factor values $(f_1, ...f_T)$. The Probability of staying in category "3" is:

$$\begin{aligned} p_{33,t} &= P[y^*_{i,t} > c_2 | f_t, y_{i,t-1} = 3], \\ &= P[\delta_3 + \beta_3 f_t + \sigma_3 u_{i,t} > c_2], \\ &= P[u_{i,t} > \frac{c_2 - \delta_3 - \beta_3 f_t}{\sigma_3}], \\ &= 1 - \Phi(\frac{c_2 - \delta_3 - \beta_3 f_t}{\sigma_3}). \end{aligned}$$

The transition probabilities from category 3 to 2 is characterized by:

$$p_{32,t} = P[c_1 < y^*_{i,t} \leq c_2 | f_t, y_{i,t-1} = 3],$$
$$= \Phi(\frac{c_2 - \delta_3 - \beta_3 f_t}{\sigma_3}) - \Phi(\frac{c_1 - \delta_3 - \beta_3 f_t}{\sigma_3}),$$

and lastly, the transition from category 3 to 1:

$$p_{31,t} = P[y^*_{i,t} \leq c_1 | f_t, y_{i,t-1} = 3],$$
$$= \Phi(\frac{c_1 - \delta_3 - \beta_3 f_t}{\sigma_3}).$$

Each row of the transition matrices is an ordered-Probit as the 3 transition probabilities are linked due to the presence of the thresholds c_1 and c_2. The ratios $(\frac{c_k - \delta_l - \beta_l f_t}{\sigma_l})$ in the above transition probabilities identify semi-parametrically the micro-parameters and the factor values up to location and scale transformations. There are identification problems due to the factor dynamics and the partial observability of the quantitative score $y^*_{i,t}$. As it stated before in this approach we consider two identification constraints for the factor dynamics and two identification restrictions for the micro-parameters, see Gagliardini and Gouriéroux (2005) for the complete discussion on the identification restrictions on the stochastic migration model. The first identification constraints is about the factor dynamics. Since, f_T is unobservable, and has to be defined up to an invertible linear transformation either additive or multiplicative:

$$f_t \to f_t + constant,$$

$$f_t \to \lambda f_t,$$

therefore, two identification constraints can be imposed for factor dynamics f_t:

$$E(f_t) = 0,$$

$$V(f_t) = 1.$$

The second identifiability problem is due to partial observability of the quantitative score. Since the same combination of affine transformations of score and threshold can give the same migration probabilities, two standard identifications $c_1 = 0$ and $\sigma_1^2 = 1$ are needed to be imposed. These constraints have to imposed to link rows of transition matrix and to allow for differences between the micro- and macro-parameters.

From the data in migration matrices, the values of $p_{33,t}$, $p_{32,t}$ and so on, are approximated by $\hat{p}_{33,t}$, $\hat{p}_{32,t}$, and so on, respectively. Under the granularity approach, it is assumed that as $n \to \infty$, Thus the cumulative probability $p^*_{lk,t}$ is well approximated by its cross-sectional sample counterpart. Therefore we can write:

$$\Phi^{-1}(p^*_{lk,t}) = \frac{c_k - \delta_l - \beta_l f_t}{\sigma_l} \text{ for } l = 1, 2, 3 \text{ and } k = 1, 2, 3.$$

Now the model can be re-parameterized in the appropriate way to obtain the canonical factors $a's$ as follows:

- First row:

$$a_{33,t} = \Phi^{-1}(p_{33,t}) = \tfrac{c_2 - \delta_3 - \beta_3 f_t}{\sigma_3},$$
$$a_{32,t} = \Phi^{-1}(p_{32,t} + p_{33,t}) = \tfrac{c_1 - \delta_3 - \beta_3 f_t}{\sigma_3},$$

- second row:

$$a_{23,t} = \Phi^{-1}(p_{23,t}) = \tfrac{c_2 - \delta_2 - \beta_2 f_t}{\sigma_2},$$
$$a_{22,t} = \Phi^{-1}(p_{22,t} + p_{23,t}) = \tfrac{c_1 - \delta_2 - \beta_2 f_t}{\sigma_2},$$

- third row:

$$a_{13,t} = \Phi^{-1}(p_{13,t}) = \tfrac{c_2 - \delta_1 - \beta_1 f_t}{\sigma_1},$$
$$a_{12,t} = \Phi^{-1}(p_{12,t} + p_{13,t}) = \tfrac{c_1 - \delta_1 - \beta_1 f_t}{\sigma_1}.$$

There are 6 approximately affine transformation of f_t, which can be written as follows:

$$a_{lk,t} = \Phi^{-1}(p^*_{lk,t}) \approx \tfrac{c_k - \delta_l}{\sigma_l} - \tfrac{\beta_l}{\sigma_l} f_t.$$

There are two alternatives ways to define the representation of factors:

- One approach is as follows: Since factors are affine linear transformation, we can choose one of the canonical factors (one cell in the transition matrix) and consider it as the common factor. In order to impose identification restrictions on the factor dynamics, we can demean and standardize it. Then, we get the following factor representation:

$$\hat{f}_t = \frac{a_{33,t} - \frac{1}{T}\sum_{t=1}^{T} a_{33,t}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(a_{33,t} - \bar{a}_{33})^2}}.$$

- An alternative approach is the following: We determine the number of factors first, by using the principal component analysis (PCA), that is the spectral decomposition of the $(T*T)$ matrix (YY'). This matrix contains the series of estimated canonical factors $\hat{a}_{l,k,t}$ where each row of that matrix is $(\hat{a}_{lk,t} - \bar{a}_{lk})$, where, $(\bar{a}_{lk} = \frac{1}{T}\sum_{t=1}^{T} a_{lk,t})$. Then, we need to find the eigenvalues and eigenvectors associated with these eigenvalues. Next, we get the number of factors from the number of the largest eigenvalues.

Under the second approach, we can impose the identification constrains on the factor dynamics by demeaning and standardizing the largest eigenvalues or the associated eigenvectors. In this case we need less constraints on the micro-parameters. Below we consider the second representation of factors, and two restrictions on factor dynamics and micro-parameters will be imposed.

Due to the identification restrictions on factor dynamics, and given $E(f_t) = 0$, we can identify:

$$\frac{c_k - \delta_l}{\sigma_l} = \frac{1}{T}\sum_{t=1}^{T} p_{lk,t},$$

$$\frac{\beta_l}{\sigma_l} = \text{Standard error of } \hat{p}_{lk,t}.$$

Next given the other identification restrictions on δ and σ, we can consistently estimate all the parameters. The complete derivation of the asymptotic properties of the estimators can be found in Gagliardini and Gourieroux (2014b), which is a comprehensive study of granularity theory in finance and insurance.

2.2.2 The Two-Step Efficient Estimation Approach

This method distinguishes two types of parameters. The first type of parameters is vector $\theta = (c_k,\ k = 1, ..., K-1, \beta_l, \sigma_l, \delta_l,\ l = 2, ..., K)$ of micro-parameters. The second set of parameters are the values of unobserved factor $f_t,\ t = 1, ..., T$. In Bayesian statistics, a model is usually seen as a function of observations given the parameters, and the parameters are considered stochastic. Thus, in our model, the vector θ, can be considered as a parameter vector, in terms of the classical statistical theory and the factor values $f_t,\ t = 1, ..., T$, are the parameters in terms of the Bayesian statistical theory. Under the granularity-based approach, the stochastic factors are also parameters. Hence, the classical methodology can be applied to estimate those parameters. Under the granularity theory, the approach is the following.

The joint density of variables $\underline{y_T} = y_{it}$ for $i = 1, ..., n,\ t = 1, ..., T$, and factor $F = (f_t)$, $t = 2, ..., T$ (conditional on the initial observations) if both individual ratings $y_{i,t}$ and f_t were observed would be as follows:

$$l^*(\underline{y_T}, F; \theta, \rho) = [\prod_{t=1}^{T} \psi(f_t|f_{t-1}; \rho)] \prod_{t=1}^{T} \prod_{i=1}^{n} h(y_{i,t}|y_{i,t-1}, f_t; \theta).$$

The log likelihood function can be defined and decomposed as:

$$\mathcal{L}^*(\underline{y_T}, F; \theta, \rho) = \log\ l^*(\underline{y_T}, F; \theta, \rho) = L^M(f_T; \rho) + \sum_{t=1}^{T} L^{CS}(\tilde{y}_t|\tilde{y}_{t-1}, f_t; \theta).$$

This log-likelihood function contains two parts, the first component $L^M(f_T; \rho)$ corresponds to the macro-log-likelihood function and the second component, $\sum_{t=1}^{T} L^{CS}(\tilde{y}_t|\tilde{y}_{t-1}, f_t; \theta)$ corresponds to the cross-sectional micro log-likelihood function:

Macro log-likelihood:

$$L^M(f_T; \rho) = \sum_{t=1}^{T} \log \psi(f_t|f_{t-1}; \rho). \tag{2.2.2.1}$$

Micro log-likelihood:

$$L^{CS}(\tilde{y}_t|\tilde{y}_{t-1}, f_t; \theta) = \sum_{i=1}^{n} \log h(y_{i,t}|y_{i,t-1}, f_t; \theta). \tag{2.2.2.2}$$

As already mentioned, the true log-likelihood function is deduced by integrating out the unobserved factor, Thus, it is given by

$$l(\underline{y_T};\theta,\rho) = \log \int ... \int \left([\prod_{t=1}^{T} \psi(f_t|f_{t-1};\rho)] \prod_{t=1}^{T} \prod_{i=1}^{n} h(y_{i,t}|y_{i,t-1},f_t;\theta) \prod_{t=1}^{T} df_t \right), \qquad (2.2.2.3)$$

Due to the complexity of equation (2.2.2.3) the likelihood function $l(\underline{y_T};\theta,\rho)$ is replaced by $\mathcal{L}^*(\underline{y_T},F;\theta,\rho)$.

If the micro-parameters θ are known the factor value at time t can be approximated by a fixed effects estimator. Thus, in the first step we get the factor values at time t as follows:

- Step 1:

$$\hat{f}_{n,t}(\theta) = \arg\max_{f_t} \sum_{i=1}^{n} \log h(y_{i,t}|y_{i,t-1},f_t;\theta). \qquad (2.2.2.4)$$

In this approach, $\hat{f}_t(\theta)$ is treated as a parameter in the latent cross-sectional micro-likelihood, and the value of the factor is estimated as a fixed time effect. In the second step, the factor approximations $\hat{f}_n, t(\theta)$ are reintroduced in the latent micro-likelihood functions, aggregated over time to obtain a function of the observations $\underline{y_T}$ and parameter θ estimated as follows:

- Step 2:

$$\hat{\theta}_{n,T} = \arg\max_{\theta} \sum_{t=1}^{T} \sum_{i=1}^{n} \log h(y_{i,t}|y_{i,t-1},\hat{f}_{n,t}(\theta);\theta). \qquad (2.2.2.5)$$

Next, an approximation of factor value at time t can be defined by reintroducing the estimated micro-parameters into the expression of the fixed effects estimator of the factor values as follows:

$$\hat{f}_{n,T,t} = \hat{f}_{n,t}(\hat{\theta}_{n,T}). \qquad (2.2.2.6)$$

After estimating T values of factors, the estimator of the macro-parameter ρ can be obtained by the maximum likelihood estimation of the autoregressive model as follows:

$$\hat{f}_{n,T,t} = \mu + \rho \hat{f}_{n,T,t-1} + \eta_t, \quad \eta_t \sim IID\ N(0,\sigma_\eta^2), \quad t = 1,...,T. \qquad (2.2.2.7)$$

In the first step, we use the estimates of micro-parameters of the approximate linear model. Thus, the micro-parameters can be considered as known parameters. Then, we estimate our model in two steps as explained above. Since the cross-sectional dimension in our framework is much larger that the time dimension, the unknown factor values can be treated as nuisance parameters without having an incidental parameter problem. This problem arises when the time dimension T tends to infinity. [see, e.g. Neyman et al. (1948), Gagliardini and Gourieroux (2014b)].

2.2.3 The Joint Optimization Approach

In the joint optimization approach, we estimate the parameters of the stochastic migration model in one step. Equivalently, the joint optimization can be performed to obtain the solution in θ and $(f_1, ..., f_T)$ for small T:

$$\max_{\theta, f_1, \ldots, f_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \log h(y_{i,t} | y_{i,t-1}, f_t; \theta). \tag{2.2.3.1}$$

In both approaches, the vectors of micro- and macro-parameters are identified after imposing the standard identification restrictions for partial observability. Since, the quantitative score is partially observable, various combinations of affine transformations of the quantitative score and of the thresholds can provide the same migration probabilities. Thus, each row of the transition matrix contains the same transformations, because, the thresholds are independent of the initial rating class. Hence, it is sufficient to impose the standard identification restrictions for an ordered-Probit model on one row only. The identification restrictions for micro-parameters in this model concern the parameters of rating class "B":

$$c_3 = \delta_4 = 0 \text{ and } \beta_4 = \sigma_4 = 1.$$

In our framework of the stochastic migration model with a single factor and the absorbing state $k = K$, it is enough to impose the identifying restrictions for partial observability only.

Gagliardini and Gourieroux (2014b) use the two-step estimation approach and derive the asymptotic properties of the micro- and macro-parameter estimators. These estimators converge with different rates of convergence to their corresponding true values when both n and T tend to infinity. The micro-parameters $\hat{\theta}_T$ are $\sqrt{nT}$ consistent and the rate of convergence is $\frac{1}{\sqrt{nT}}$. The macro-parameter $\hat{\rho}_T$ is $\sqrt{T}$ consistent and the rate of convergence is $\frac{1}{\sqrt{T}}$. The factor estimates $\hat{f}_t$ are $\sqrt{n}$ consistent and the rate of convergence is $\frac{1}{\sqrt{n}}$. The estimators are asymptotically efficient and they are asymptotically equivalent to the maximum likelihood estimators. In the case when both n and T are large, the lowest possible variance can be obtained for the estimators $\hat{\theta}_{n,T}$ and $\hat{\rho}_T$. [see, Gagliardini and Gourieroux (2014b) for the further discussion on the asymptotic properties of the micro- and macro parameters]

2.3 Data

The credit ratings are the key factor for credit risks evaluation. Past changes in credit quality of obligors are characterized by the credit migration or transition matrices. These matrices provide information about the probabilities of change in ratings in one year. The transition matrices are regularly reported by the rating agencies such as the Moody's and Standard and Poor's which are external rating systems, or can be reported by banks which have their own internal rating systems. In the credit risk management, using the internal credit ratings is becoming more and more frequent due to Basel II requirements for implementation of an internal rating approach (IRB) [Basel Committee on Banking Supervision (2001, 2003)]. In order to understand how banks determine their lending function and how they control their risk exposures, it is essential to study how rating systems are designed, and used in risk management.

A bank must choose the construction of its ratings system and must decide which loss measure such as default, recovery, exposure, and expected loss to employ. Each of these loss measures corresponds to a grade on the ratings scale. Treacy and Carey (1998) showed that the 50 largest banks in US use two kinds of rating systems which are one and two-dimensional. Among these banks, almost 60% have the one-dimensional rating system and 40% have the two-dimensional rating system. In another words, 60% of banks, approximate the expected loss only and remaining 40% assess the default risk of obligors on one scale, and the expected loss of individual exposure on another scale. Typically, the number or letters are used to express the rating scale. For instance, in the external rating system with letters, investment grades are represented by AAA, AA, A and BBB which are the four highest categories respectively. Grade BB and below represent the non-investment grade. Most of the rating agencies (external rating systems) prefer to provide rankings of large firms at the international level while banks (internal rating systems) use the balance sheets of all firms to construct a quantitative score that evaluates the default probabilities of firms depending on a set of financial ratios and individual characteristics.

2.3.1 Description of the Data

The data set consists of 8 (eight) migration matrices that contain the internal rating data provided by Credit Agricole S.A. bank with headquarter in Montrouge, France. These matrices were recorded annually from 2007 to 2015, hence $T = 8$ in our sample. There are 14 risk categories in the Credit Agricole S.A. credit rating system. These categories are "A+", "A", "B+", "B", "C+", "C", "C-", "D+", "D", "D-", "E+", "E", "E-" and "F", from the lowest to the highest risk. All rating movements over a one year period are reported in transition matrices. These movements are migrations from one rating to another, except for default category which is the absorbing state and once a company joins the default category, there is no exit from it. Table 2.3 shows an example of a transition matrix for year 2007, which has 14 categories of risks. The rating categories from which transitions are to be made are represented in rows. The first row refers to rating category "A+", and the last row shows the category "D". However, the category "F", is excluded from the rows, since no company can exit from the default category. Columns 1 to 14 correspond to ratings categories "A+" to "F" to which transitions will be made until the end of year 2007. The numbers of long-term rated issuers per rating category at the beginning of the year 2007 are reported in the issuer column. Column 15, "NR", relates to issuers, who were rated at the beginning of year 2007 but were not rated at the end of year 2007. From the statistical point of view, there are reasons that cause this lack of information. For example, sometimes a rating cannot be assigned to a firm. For instance, when the debt of a firm is completely paid off or when a firm is terminated and the relevant debt extinguished, the rating of that firm is eliminated [see Brady et al. (2002)]. Each element in Table 2.3 represents the observed transition frequencies for year 2007. For instance, the second row shows that, out of "322" firms rated "A" at the beginning of year 2007, 2.8 % were upgraded to "A+", 27.33% stayed in the same rating category, and 24.53% were downgraded to rating category "B+". Typically, the highest percentages are found on the main diagonal of the transition matrix, because most of the firms remain in the same credit ratings category. The last column represents the proportion of non-rated corporations.

Table 2.3: Number of Issuers and Transition Matrix for Year 2007 in %

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	2007	ISSUER	A+	A	B+	B	C+	C	C-	D+	D	D-	E+	E	E-	F	NR
1	A+	91	53	13	13	10	3	2	4	0	0	0	1	0	0	0	1
2	A	322	2.8	27.33	24.53	11.8	7.76	9.63	3.42	5.59	3.42	1.24	0.31	1.24	0.62	0.31	0
3	B+	1132	0.88	2.39	28.62	11.93	13.34	16.43	9.36	6.45	4.51	3.36	1.41	0.62	0.62	0.09	0
4	B	3181	0.03	0.16	2.77	23.45	14.4	19.96	15.72	11.76	5.94	2.92	1.54	0.53	0.5	0.31	0.01
5	C+	2904	0.07	0.03	2.07	4.75	34.92	13.5	15.5	12.36	8.54	3.99	2.1	0.93	0.55	0.69	0
6	C	8120	0.01	0.1	0.83	1.98	7.46	29.31	17.4	19.47	12.4	5.99	2.46	0.92	0.84	0.83	0
7	C-	7263	0.04	0.1	0.54	0.81	3.95	6.83	36.09	18.24	18.08	9.14	3.37	1.25	0.77	0.78	0.01
8	D+	6252	0.03	0.03	0.22	0.75	2.91	5.52	11.72	42.24	17.93	11.28	3.82	1.6	0.86	1.07	0.02
9	D	14256	0.02	0.04	0.14	0.32	1.36	3.31	9.48	13.73	38.53	20.19	6.76	2.62	1.98	1.52	0
10	D-	9207	0.01	0.02	0.09	0.24	1.08	1.8	6.1	9.51	15.9	43.04	10.49	4.95	3.65	3.11	0.01
11	E+	3356	0.03	0	0.12	0.06	0.6	1.4	3.81	5.9	11.86	19.85	34.51	9.71	6.94	5.21	0
12	E	2641	0.04	0	0.08	0.38	0.27	1.1	3.1	5.26	9.01	16.74	13.86	31.58	11.66	6.93	0
13	E-	2451	0.08	0.12	0	0.16	0.69	1.43	1.55	3.59	6	9.63	9.47	9.71	48.92	8.65	0

In the literature, the following two alternative approaches were considered, which are:

- Include the information on transition probabilities of non-rated firms at the beginning of the year
- Exclude the information on transition probabilities of non rated firms at the end of the year

The first approach requires information on companies that are not rated at the beginning of each year. However, since this information is not easily available, the second option is often considered. Nickell et al. (2000), Bangia et al. (2002) and Foulcher et al. (2005), have all excluded the information on "NR" firms from the transition matrices and they allocated the weights of "NR" firms among the other risk categories. In this paper, the same approach is followed. Table 2.4 shows the matrix of migration probabilities from Table 2.3 adjusted for N.R. For instance, in the first row of the "N.R.-adjusted" transition matrix the frequency of transition from "A+" to "A+" is 53.53% which is calculated from the ratio $53/(1-0.01)$.

Table 2.4: Adjusted Transition Matrix for Year 2007

	A+	A	B+	B	C+	C	C-	D+	D	D-	E+	E	E-	F
A+	53.54	13.13	13.13	10.10	3.03	2.02	4.04	0.00	0.00	0.00	1.01	0.00	0.00	0.00
A	2.80	27.33	24.53	11.80	7.76	9.63	3.42	5.59	3.42	1.24	0.31	1.24	0.62	0.31
B+	0.88	2.39	28.62	11.93	13.34	16.43	9.36	6.45	4.51	3.36	1.41	0.62	0.62	0.09
B	0.03	0.16	2.77	23.45	14.40	19.96	15.72	11.76	5.94	2.92	1.54	0.53	0.50	0.31
C+	0.07	0.03	2.07	4.75	34.92	13.50	15.50	12.36	8.54	3.99	2.10	0.93	0.55	0.69
C	0.01	0.10	0.83	1.98	7.46	29.31	17.40	19.47	12.40	5.99	2.46	0.92	0.84	0.83
C-	0.04	0.10	0.54	0.81	3.95	6.83	36.09	18.24	18.08	9.14	3.37	1.25	0.77	0.78
D+	0.03	0.03	0.22	0.75	2.91	5.52	11.72	42.25	17.93	11.28	3.82	1.60	0.86	1.07
D	0.02	0.04	0.14	0.32	1.36	3.31	9.48	13.73	38.53	20.19	6.76	2.62	1.98	1.52
D-	0.01	0.02	0.09	0.24	1.08	1.80	6.10	9.51	15.90	43.04	10.49	4.95	3.65	3.11
E+	0.03	0.00	0.12	0.06	0.60	1.40	3.81	5.90	11.86	19.85	34.51	9.71	6.94	5.21
E	0.04	0.00	0.08	0.38	0.27	1.10	3.10	5.26	9.01	16.74	13.86	31.58	11.66	6.93
E-	0.08	0.12	0.00	0.16	0.69	1.43	1.55	3.59	6.00	9.63	9.47	9.71	48.92	8.65

In addition, for computational simplicity, the rating categories are reduced from 14 to 7 categories. These categories are ranked from the lowest to the highest risk. In this paper, for computational simplicity, these 14 categories are reduced to 7 categories as follows. We keep the two credit ratings "A+" and "A" that are considered high grades or investment groups and the two categories "B+", "B" that are considered medium grades or non-investment groups. The categories of "C+", "C" and "C-" are aggregated and become one category "C". The remaining rankings except for category "F" are aggregated to "D". The ratings "C" and "D" are considered low grades or substantial risks groups and "F" stands for default. For computational convenience this rating scheme is replaced by quantitative indicators, "1", "2", ..., "7". As shown in Table 2.5, there are 7 states among, which state "1" or default, "F", for the lowest rating category or the highest risk firms, which is a terminal rating of a firm and "7" represents the the highest rating category or the lowest risk firms, "A+". Table 2.5 below, shows the rating scheme that summarizes the approach. The firms are assigned to a given ranking depending on their score being less or greater than a given threshold $c_i,\ i = 1, ..., 6$ at the beginning of each year. Accordingly, the rankings of a firm change over time.

Table 2.5: Rating Scheme

R.C		TH
7	A+	$\leq c_6$
6	A	$c_5 \leq ...< c_6$
5	B+	$c_4 \leq ...< c_5$
4	B	$c_3 \leq ...< c_4$
3	C	$c_2 \leq ...< c_3$
2	D	$c_1 \leq ...< c_2$
1	F	$< c_1$

The new adjusted transition matrix, after aggregating the three categories of "C's" to one category "C" and after aggregating the categories "D's" and "E's" to one category "D", is given in Table 2.6:

Table 2.6: Adjusted Transition Matrix for Year 2007

2007	Issuers	A+	A	B+	B	C	D	F
A+	91	53.54	13.13	13.13	10.10	9.09	1.01	0.00
A	322	2.80	27.33	24.53	11.80	20.81	12.42	0.31
B+	1132	0.88	2.39	28.62	11.93	39.13	16.97	0.09
B	3181	0.03	0.16	2.77	23.45	50.09	23.19	0.31
C	18287	0.03	0.22	0.91	1.96	52.82	43.40	0.79
D	38163	0.03	0.03	0.13	0.34	11.81	84.68	2.99

The aggregated adjusted transition matrix keeps the structure of the original migration matrix. However, due to the aggregation, we observe that the probabilities of downgrades from categories "B+", "B", "C" and "D" to categories "C" and "D" are higher than, or close to the stability rates of these ratings over the active sampling period. For instance, in Table 2.6 the frequency of transition from "B+" to "B+" is 28.62% but the frequency of downgrade from rating "B+" to rating "C" is 39.13% which is higher than stability rate of rating "B+". In the data set, we observe that pattern in years 2007 and 2011 which proceeded the consecutive recessions in France when many firms were downgraded to lower categories. The two diagonals above and below the main diagonal show transition probabilities are relatively high, while the other entries are close to zero. The frequencies of default are given in the last column of the migration matrix. As shown in the Table 2.6, the probabilities of default increase monotonically with a deteriorating credit quality. This is true for each of 8 migration matrices in the data set. The original and aggregated adjusted transition matrices for the whole sampling period, 2007-2014, are reported in Appendix A.1 and A.2 respectively.

The number and the distribution of firms across the rating classes has changed over the sampling period in 2007 to 2014. Table 2.7 reveals that the total number of firms increased by 21% from $63,183$ to $76,589$ firms. The class "D" contains both category "D's" and "E's", which are relatively poor-quality firms. This change can be due to the fact that the average credit quality of the firms has decreased, or the rating system of the bank has become more strict.

Table 2.7: Distribution of Firms Across Rating Classes in 2007 and in 2014

	A+	A	B+	B	C	D	Total
2007	91	322	1132	3181	18287	38163	63183
2014	10	105	410	1005	14670	58375	76589

In the next step, let us examine the dynamics of the transition matrices. We consider four time series i.e. the time series of frequencies of 1) staying in the same class (called the stability rate, henceforth), 2) upgrade to higher class, 3) downgrade to lower category and 4) the time series of rate of default. We also display the average default rate of each rating categories.

Stability Rate of Risk Categories

The main diagonal of each transition matrix indicates the frequency of staying in the same class of ratings. Figure 2.1 presents the time series of stability rates for all rating classes.

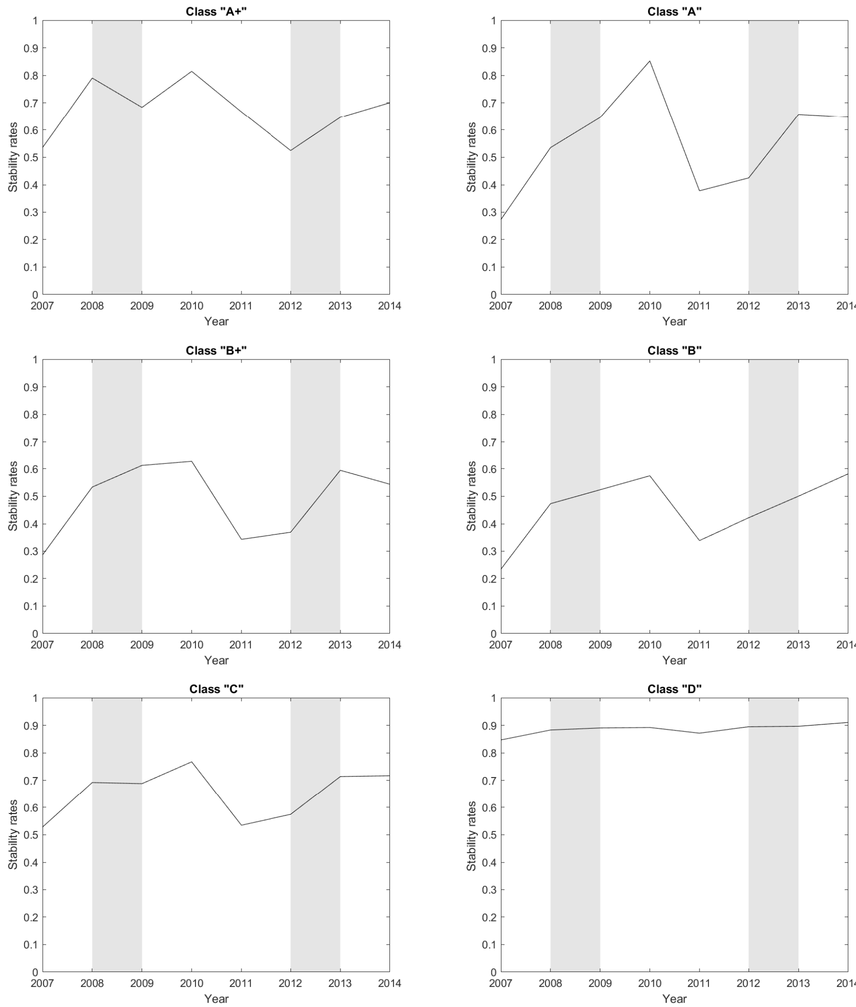


Figure 2.1: Stability Rates. This figure shows the time series of probabilities of staying in the same class. The migration rates are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

As shown in Figure 2.1, the probability of staying in the rating class "A+" varies between 50-82%, while for the rating classes "A" and "B+", it varies between 28-86% and 29-63% respectively. All rating classes show an upward trend in their stability rates before the economic recession in France in 2008-2009 an a downward trend in their stability rates before the economic recession in France and the Euro area in 2011-2013 as documented by the OECD.

Downgrade and Upgrade Probabilities

The numbers to the left and right of the main diagonal of a transition matrix represent the upgrading and downgrading rates respectively. For instance, the downgrade rates are defined as: $d_{l,t} = \sum_{k=1}^{l+1} \hat{p}_{lk,t}$ and the upgrade rates are defined as: $u_{l,t} = \sum_{k=1}^{l-1} \hat{p}_{lk,t}$. Figure 2.2 shows the time series of downgrades for each rating class. Figure 2.3 presents the upgrade rates for each rating class. We observe that the migration rates are changing over time.

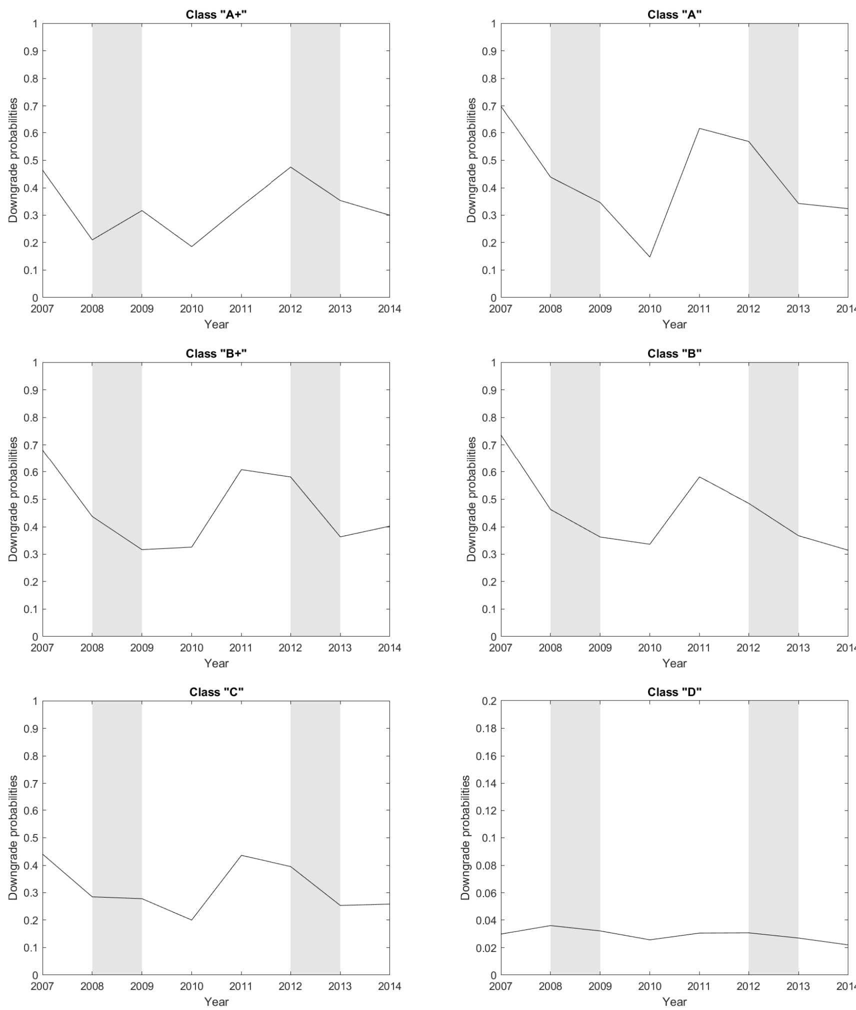


Figure 2.2: Downgrade Rates. This figure shows the time series of frequencies of downgrade for each rating class. The migration frequencies are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

In Figure 2.2, rating category “A+” reveals an overall increase in downgrade rates during the economic recession periods of 2008-2009 in France. All rating categories show an increase in downgrade rates before the economic recession in 2012-2013 in France. Figure 2.3 demonstrates the upgrade rates for each rating categories. For instance, rating class “A+” shows downward trends in upgrade rates during the economic recessions 2008-2009 and 2012-2013 in France.

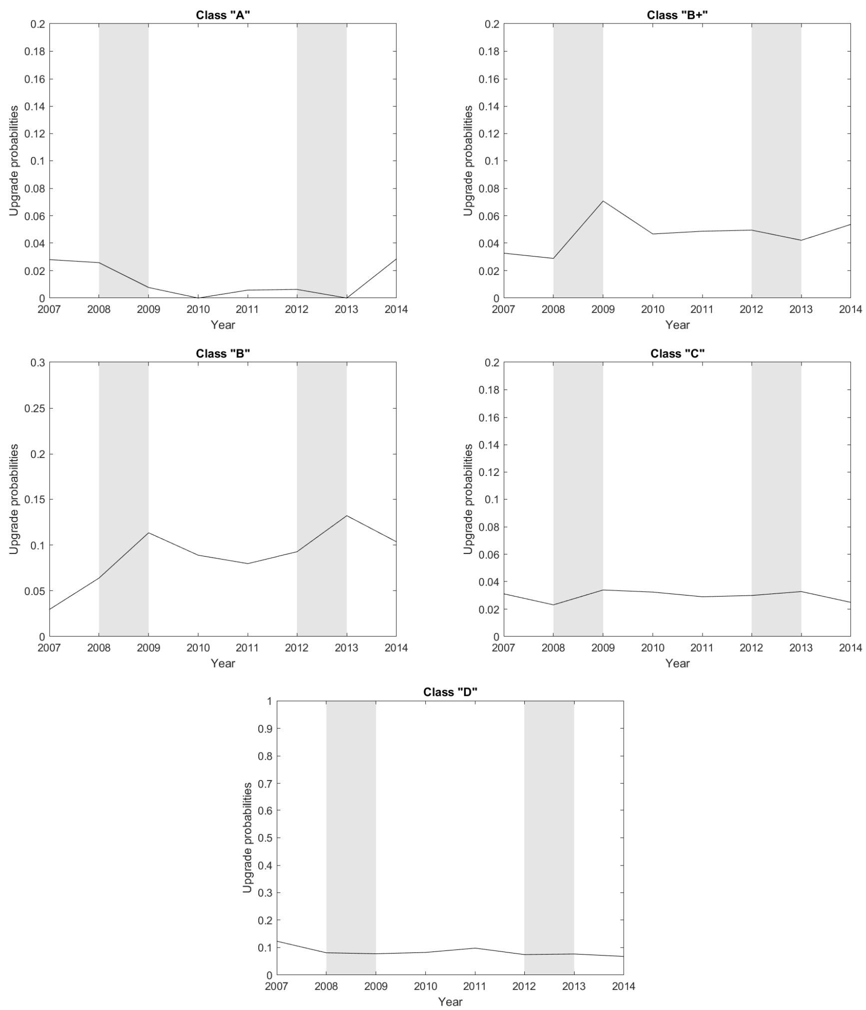


Figure 2.3: Upgrade Rates. This figure shows the time series of frequencies of upgrade for each rating class. The migration rates are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

Probability of Default

Figure 2.4 displays the time series of rates of default from rating categories “B+”, “B”, “C” and “D”. We plot the time series of default rates for rating class “B+”, “B”, “C” and “D” only, because the “A+” and “A” categories have null default rates. We observe that the rates of default for rating categories “B+”, “B”, “C” and “D” have an upward trend before the economic recession in 2008-2009 in France. Moreover, these rating categories have an downward trend before the other economic recessions in France in 2012-2013.

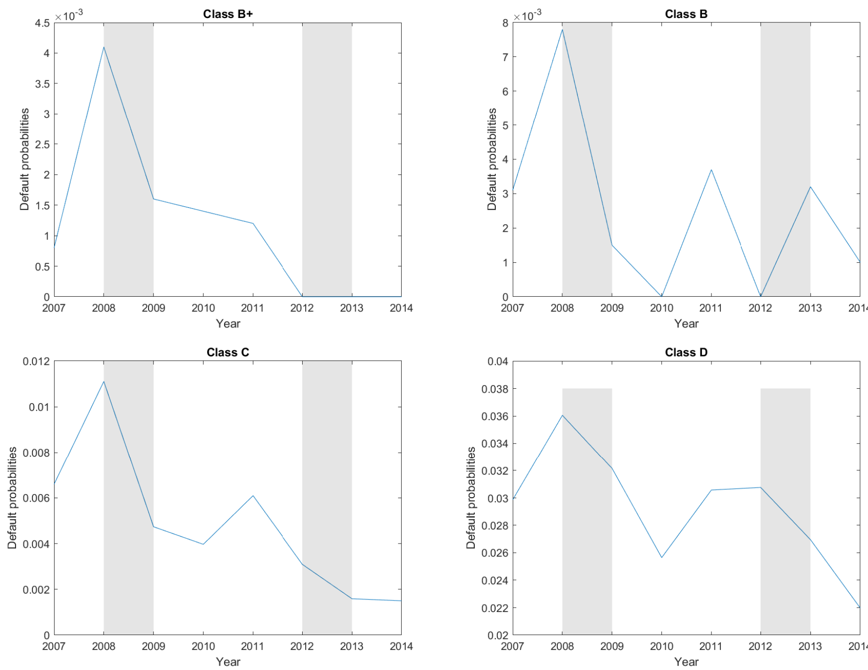


Figure 2.4: Probability of Default. This figure shows the time series of probabilities of default for rating class “B+”, “B”, “C” and “D”. The default probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

Figure 2.5, displays the average default rates of each rating category. The horizontal axis shows the rating classes. Rating 7 refers to rating class “A+” and rating 2 corresponds to rating class “D”. The average default rate increases as the rating class becomes more riskier.

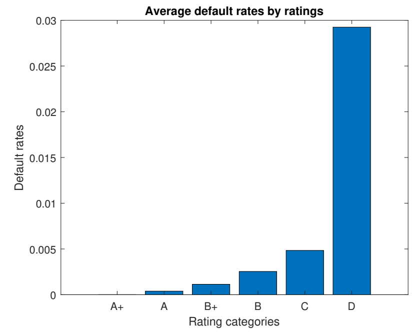


Figure 2.5: Average Default Rates by Ratings.This figure shows the average frequencies of default for each rating class. The default rates are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

In the next section, the transition matrices for $t = 1, ..., T$ are used in the estimation of the stochastic factor ordered-Probit model with $k = 7$ rating classes.

2.4 Estimation Results

This section presents the estimation of the model described in Chapter 1, Section 1.5.1, using the granularity-based maximum likelihood methods discussed in Section 2.2 of this Chapter. In Section (2.4.1), we provide the estimation results on the linearly approximated state-space model that is used as a benchmark model to provide parameter values for the two-step estimation procedure. In Section (2.4.2), we provide the two-step estimation results. The results of the joint optimization approach are presented in the Section (2.4.3).

2.4.1 Benchmark Model-Linear Approximation of State-Space Model

Let us write the ordered-Probit stochastic transition model as an approximate linear state-space model and apply the procedure explained in Section 2.2.1. As explained earlier, before estimating the parameters we need to determine how many latent factors are driving the model. To do so, the principal component analysis (PCA), has to be performed. The PCA is the spectral decomposition of square $(T \times T)$ matrix of the series of canonical factors $\hat{a}_{lk,t}$ and it is based on the analysis of eigenvalues and eigenvectors of this matrix. The PCA converts a set of correlated observed variables into a set of linearly uncorrelated variables by using orthogonal transformations. This transformation, provides the eigenvectors corresponding to the eigenvalues of the variance-covariance matrix arranged in an ascending order. In other words, the first principal component, is the one associated with largest eigenvalue. The motivation for doing this is to find the number of factors that have caused changes in migration probabilities across the rating classes. Therefore, the first objective is to compute the canonical factors $\hat{a}_{lk,t}$ from the observed frequencies $\hat{p}_{lk,t}$. The next step is to define the square matrix of series of these canonical factors which has the dimension $(T \times T)$ too. Then, using the principal component analysis (PCA) we can sort the eigenvalues in decreasing order and determine the associated eigenvectors. The ordered qualitative stochastic transition model with $K = 7$ is estimated as follows:

First, the model is re-parametrized to compute the canonical factors $\hat{a}_{lk,t}$ from the observed frequencies $\hat{p}_{lk,t}$. There are eight years ($T = 8$) of transition probabilities matrices. Next, the cumulative probabilities of these frequencies are computed. Then the quantile function of a standard normal distribution is applied to obtain the canonical factors. In the next step, the matrix (YY'), which is the $(T \times T)$ matrix of series of estimated canonical factors is computed and the principal component analysis (PCA) is performed. These eigenvalues are reported in Table 2.8:

Table 2.8: The Eigenvalues

10.4834	2.6814	1.1795	0.7452	0.4709	0.4003	0.1393	-5.6669e-16

These eigenvalues are derived from the spectral decomposition of the sample variance-covariance matrix of the series $\hat{a}_{l,k,t}$. The difference between the largest eigenvalue and the next one is significantly greater than the remaining differences. Moreover, we observe that the first eigenvalue is much larger than the other ones. Therefore, there is a cut-off in the sequence of eigenvalues. This evidence based on the PCA analysis suggests us to consider the stochastic model with one-factor. It means that, each year there is a common factor for all rating classes that drives the migration probabilities. The normalized and standardized eigenvectors associated with the eigenvalues in Table 2.8 are reported in Table 2.9.

Table 2.9: The Eigenvectors Associated with the Eigenvalues

-5.67E-16	0.139355	0.400382	0.470923	0.745274	1.179573	2.681447	**10.4835**
0.3455	0.5349	-0.2112	0.8822	-0.4852	-1.4001	-0.2771	**−1.6564**
-0.1728	-0.5822	1.1272	0.1427	1.9515	0.0049	-0.8115	**−0.1675**
0.8207	0.3996	0.0517	-2.1490	-0.2476	-1.0096	-0.2125	**0.4676**
-0.4751	-0.5762	-0.9726	0.6944	-0.6699	0.1513	-1.5887	**1.1702**
1.0798	0.7623	-1.2716	-0.5046	0.5223	1.6188	0.1224	**−0.8740**
-2.1597	-1.6246	0.8059	-0.2937	-1.1804	0.7514	0.6109	**−0.6471**
0.3024	-0.4749	-0.8507	0.5598	0.6767	-0.6931	1.7790	**0.8714**
0.1728	1.5611	1.3214	0.6683	-0.5674	0.5764	0.3774	**0.8357**

The pattern of factor corresponding to largest eigenvalue is shown in Figure 2.6.

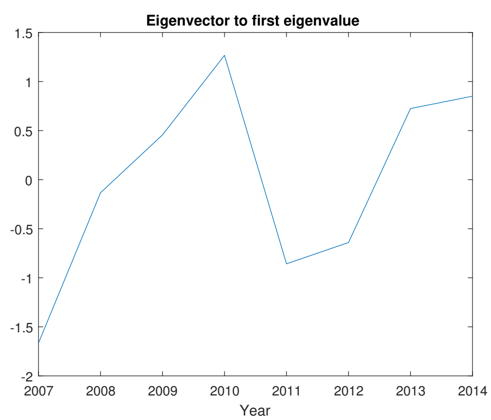


Figure 2.6: Eigenvector Obtained by the Principal Component Analysis. The figure displays the pattern of the eigenvector associated with the largest eigenvalue in the PCA of the estimated canonical factors.

The pattern of the eigenvector associated with the largest eigenvalue is consistent with the evolution of migration probabilities shown in Figures 1, 2, 3 and 4. The factor path has two peaks in years 2010 and 2013 and two troughs in years 2007 and 2011 which coincide with the peaks in stability rates and troughs in downgrade rates. Moreover, the factor path features a downward local trend in 2010-2011 corresponding to decreases in the downgrade risk and an upward local trend in years 2007-2010 corresponding to an increase in the downgrade risk.

The estimates of micro-parameters are displayed in Table 2.10.

$\hat{c}_1 = -2.434$	$\hat{c}_2 = -1.333$	$\hat{c}_3 = -0.111$	$\hat{c}_4 = 1.384$	$\hat{c}_5 = 2.345$	$\hat{c}_6 = 2.669$

$\hat{\delta}$		$\hat{\beta}$		$\hat{\sigma}$	
1	$\hat{\delta}_2 = -2.553$	1	$\hat{\beta}_2 = 0.054$	1	$\hat{\sigma}_2 = 0.880$
2	$\hat{\delta}_3 = -1.085$	2	$\hat{\beta}_3 = 0.123$	2	$\hat{\sigma}_3 = 0.510$
3	$\hat{\delta}_4 = 0$	3	$\hat{\beta}_4 = 0.353$	3	$\hat{\sigma}_4 = 1$
4	$\hat{\delta}_5 = 0.715$	4	$\hat{\beta}_5 = 0.451$	4	$\hat{\sigma}_5 = 1.336$
5	$\hat{\delta}_6 = 3.916$	5	$\hat{\beta}_6 = 1.931$	5	$\hat{\sigma}_6 = 4.111$
6	$\hat{\delta}_7 = 5.797$	6	$\hat{\beta}_7 = 1.894$	6	$\hat{\sigma}_7 = 6.838$

Table 2.10: Micro-Parameter Estimates-Linear Approximation. Thresholds $\hat{c}_k$, $k = 1, ..., 6$ intercepts $\hat{\delta}_l$, factor sensitivities $\hat{\beta}_l$, and volatilities $\hat{\sigma}_l$, $l = 2, ..., 7$ are displayed.

The thresholds are increasing functions of the rating class indexes. The estimated intercepts, factor sensitivities and volatilities, reported in the middle panel for rows $l = 2, ..., 7$ of the transition matrix, correspond to rating class “A+”, “A”, ..., “C”. As the rating category increases, the intercepts increase too. It confirms that the rating class “A+”, which is the least risky class has the score function larger than other rating classes. The estimated factor sensitivities are all positive which means that when the factor increases, the underlying scores for each rating class are improved. Lastly, the least risky rating “A+” has the highest volatility.

The estimate of macro-parameter ρ is found by replacing the factor proxies in the macro-dynamics and applying the ordinary least square estimation which is equivalent to the maximum likelihood estimation of the autoregressive process:

$$\hat{f}_{n,T,t} = \mu + \rho \hat{f}_{n,T,t-1} + \eta_t,$$

$$\eta_t \sim IID\ N(0,1) \text{ for } t = 1, ..., 8$$

Table 2.11, reports the estimated macro-coefficient. We observe that, the estimates of μ and ρ are not statistically significant.

Table 2.11: Estimated Macro-Coefficients

	Estimate	SE	T.Stat	P-Value
μ	-0.1593	0.4378	-0.3637	0.7309
ρ	0.1685	0.5615	0.3001	0.7761

Next, we discuss the results of the two-steps estimation method for the stochastic factor ordered-Probit model introduced in Section 4 with $k = 7$ rating classes. This method relies on a simplified log-likelihood. We use the estimates of the linear approximation as a benchmark model that provides initial values in the numerical algorithm to maximize the log-likelihoods in the two-steps estimation method and the joint optimization method.

2.4.2 Two-Step Efficient Estimators Approach

In the first step, we maximize the latent micro-likelihood functions L^{CS} given by equation (2.2.9) with respect to the factor values given that θ_B (the micro-parameters obtained by the linear approximation state-space model) is known as follows:

- Step 1:

$$\hat{f}_{n,t}(\theta) = \arg\max_{f_t} \sum_{i=1}^{n} \log h(y_{i,t}|y_{i,t-1}, f_t; \hat{\theta}_B).$$

In the second step, these solutions are reintroduced in the latent micro-likelihood functions L^{CS} and aggregated over time to obtain the estimator of θ as follows:

- Step 2:

$$\hat{\theta}^*_{n,T} = \arg\max_{\theta} \sum_{t=1}^{T} \sum_{i=1}^{n} \log h(y_{i,t}|y_{i,t-1}, \hat{f}_{n,t}(\theta); \theta).$$

Next, we reintroduce the estimates of $\hat{\theta}^*_{n,T}$ into the latent micro-likelihood functions to obtain the factor values which are used as proxies for the unobserved factor values as follows:

$$\hat{f}_{n,T,t} = \hat{f}_{n,t}(\hat{\theta}^*_{n,T}).$$

The estimator of the macro-parameter ρ can be derived simply by applying the maximum likelihood estimation of the autoregressive model as follows:

$$\hat{f}_{n,T,t} = \mu + \rho \hat{f}_{n,T,t-1} + \eta_t, \quad \eta_t \sim IID\ N(0, \sigma_\eta^2), \quad t = 1, ..., T.$$

This autoregressive model can also be estimated by the ordinary least square (OLS) estimator which coincides with the maximum likelihood estimator, under the normality assumption on $\eta_t,\ t = 1, ..., T$.

We impose the identification restrictions for rating class "B" as follows:

$$c_3 = \delta_4 = 0, \text{ and } \beta_4 = \sigma_4 = 1.$$

Table 2.12 displays the estimates of the micro-parameters which are all statistically significant, except for the factor sensitivity β_2 of rating category "D". The upper panel shows the estimates of the thresholds parameters c_k. As expected, they increase with the rating class indices. Moreover, we can see that the thresholds decrease with the rating quality. The estimates for the parameters in rows $l = 2, ..., 7$ of the transition matrix , which correspond to rating classes "D", "C", ..., "A+" are displayed in the lower panel.

$\hat{c}_1 = -5.141^{***}$ (0.026)	$\hat{c}_2 = -1.840^{***}$ (0.011)	$\hat{c}_3 = 0$	$\hat{c}_4 = 1.032^{***}$ (0.006)	$\hat{c}_5 = 2.143^{***}$ (0.014)	$\hat{c}_6 = 3.350^{***}$ (0.023)

$\hat{\delta}$		$\hat{\beta}$		$\hat{\sigma}$	
1	$\hat{\delta}_2 = -3.213^{***}$ (0.055)	1	$\hat{\beta}_2 = -0.011$ (0.022)	1	$\hat{\sigma}_2 = 1.033^{***}$ (0.011)
2	$\hat{\delta}_3 = -1.752^{***}$ (0.010)	2	$\hat{\beta}_3 = 0.223^{***}$ (0.001)	2	$\hat{\sigma}_3 = 0.924^{***}$ (0.005)
3	$\hat{\delta}_4 = 0$	3	$\hat{\beta}_4 = 1$	3	$\hat{\sigma}_4 = 1$
4	$\hat{\delta}_5 = 0.633^{***}$ (0.009)	4	$\hat{\beta}_5 = 0.089^{***}$ (0.002)	4	$\hat{\sigma}_5 = 1.380^{***}$ (0.009)
5	$\hat{\delta}_6 = 1.570^{***}$ (0.016)	5	$\hat{\beta}_6 = 0.189^{***}$ (0.002)	5	$\hat{\sigma}_6 = 1.564^{***}$ (0.011)
6	$\hat{\delta}_7 = 10.925^{***}$ (0.055)	6	$\hat{\beta}_7 = 1.124^{***}$ (0.005)	6	$\hat{\sigma}_7 = 10.230^{***}$ (0.059)

Table 2.12: Micro-parameter Estimates-Two-step Efficient Estimation Method. thresholds $\hat{c}_k$, $k = 1, ..., 6$, intercepts $\hat{\delta}_l$, factor sensitivities $\hat{\beta}_l$, and volatilities $\hat{\sigma}_l$, $l = 2, ..., 7$. Standard errors are given in parentheses. $*, p < 0.05$, $**$, $p < 0.01, ***$, $p < 0.001$

All the estimates are statistically significant except for the sensitivity β_2 of the rating category "D" to the common factor. The intercepts δ_l are increasing with respect to the rating index, and increasing with respect to rating quality which confirms that the underlying quantitative score $y^*_{i,t}$ for credit quality is larger for the less risky rating classes. It also shows that the downgrade risk is higher for lower credit ratings. All the sensitivities of the rating categories β_l to the common factor are positive except for β_2. This shows that an increase in the factor increases the underlying quantitative score for credit quality in all rating classes. Generally, the riskier rating categories have smaller idiosyncratic volatility σ_l.

The approximated factor values are given in Table 2.13. In order to obtain mean zero and unit variance, these values have been standardized.

Table 2.13: Approximated Factor Values

Estimated factor values							
2007	2008	2009	2010	2011	2012	2013	2014
-1.8238	-0.0424	0.5360	1.0450	-0.9329	-0.3803	0.7928	0.8057

Table 2.14 shows the estimated macro-parameter $\hat{\rho}$ which is obtained by the Maximum Likelihood estimation of the autoregression equation (2.2.12). The estimator is based on the cross-sectional approximations of the factor values $\hat{f}^*_{n,T,t} = \hat{f}_{n,t}(\hat{\theta}^*_{n,T})$.

Table 2.14: Estimated Macro-Coefficients

	Estimate	SE	T.Stat	P-Value
μ	0.2180	0.2918	0.7471	0.4549
ρ	0.4097	0.3027	1.3534	0.1759
σ^2_η	0.2740	0.3880	0.7060	0.4801

The estimator of parameter σ^2_η is given by $\hat{\sigma}^2_\eta = \frac{1}{T-1}\sum_{t=2}^{T}\hat{\eta}^2_t$, where $\hat{\eta}_t = \hat{f}_{n,T,t} - \hat{\mu} - \hat{\rho}\hat{f}_{n,T,t-1}$ are the residuals. The autoregressive coefficient is positive but not statistically significant. The approximated factor values $\hat{f}_{n,T,t}$ are displayed in Figure 2.7.

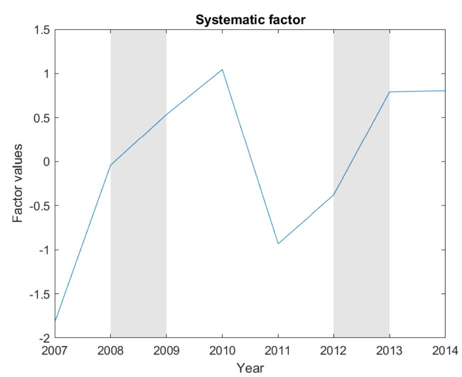


Figure 2.7: Systematic Factor. The figure displays the pattern of the systematic factor $\hat{f}_{n,T,t}$ for $t = 1, ..., 8$. The factor estimates are standardized to obtain zero-mean and unit variance in the sample. The shaded areas indicate the OECD recessions in France.

The systematic factor values $\hat{f}_{n,T,t}$ and the eigenvector associated with the largest eigenvalue obtained from the PCA analysis are very close. Figure 2.8 displays the eigenvector associated with largest eigenvalue and the systematic factor $\hat{f}_{n,T,t}$.

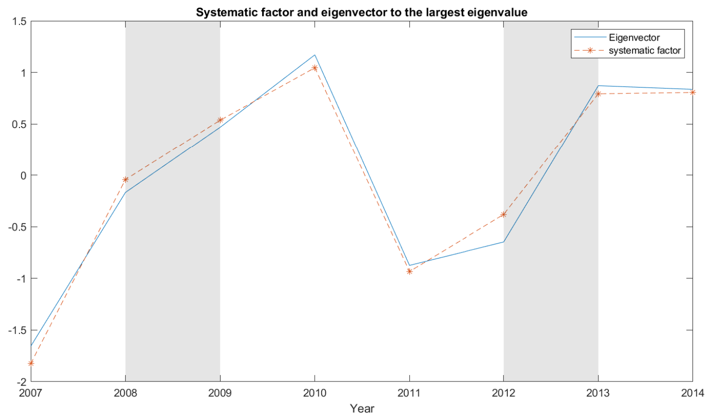


Figure 2.8: Systematic Factor and Eigenvector of the Largest Eigenvalue. This figure displays the pattern of the approximated factor values $\hat{f}_{n,T,t}$ for $t = 1, ..., 8$ (dashed line) and the eigenvector associated with the largest eigenvalue (solid line). The factor values and the eigenvector are standardized to obtain mean zero and unit variance in the sample. The shaded areas indicate the OECD recessions in France.

The cyclical pattern of the systematic factor in Figure 2.7 is consistent with the evolution of migration probabilities shown in Figures 2.1 and 2.2. As we stated before, the troughs in the factor pattern are associated with the troughs in stability rates, the peaks in downgrade rates in Figures 1,2. We also have lower values for the factor in 2007 and 2011 compares to the values of the eigenvector in 2007 and 2011. Moreover, the factor path features a downward local trend in 2010-2011 corresponding to decreases in the downgrade risk and an upward local trend in years 2007-2010 corresponding to an increase in the downgrade risk.

The next section discusses the link between the approximated factor values with the business cycle literature. The existing literature usually compares the common factor with the underlying state of economy to interpret the factor. It is very common to link the approximated factor path with the macro variables such as the business cycle, inflation rate, industrial production, unemployment or interest rate. There exists several studies which relate the failure rate with proxies of the U.S. business cycle, see e.g. Nickell et al. (2000), Bangia et al. (2002) and Rösch (2005).

2.4.3 Macroeconomic Covariates

In order to find a link between the systematic factor f_t and the economic variables, we group the macroeconomic variables in four categories of, national accounts, financial markets, stock markets and composite indices. We include variables that have been explored in previous research [see Kim (1999), Figlewski et al. (2012)]. The corporate credit risk is influenced by a variety of broad economic conditions. Thus, we select several key economic indicators to test and find the link between our systematic factor and macroeconomic variables.

National Accounts

Real GDP growth: The time series of real GDP is obtained from the World Bank. The change in GDP refers to economic strength. A higher GDP growth might influence the firm's income and a lower GDP growth increases the possibility of default for firms.

Output gap: The time series of the output gap for France is estimated by the St. Louis Federal Reserve. Although the average of the series is negative but (algebraically), stronger economic conditions cause a higher value of the output gap.

Industrial production: We include the Production of Total Industry as a better measure for the corporate sector since the GDP comprises all economic activity, including government, non-corporate business, and other sectors.

Unemployment and employment rate: The most visible measures of the overall health of the economy are the unemployment and employment rates. A low unemployment should decrease the probability of downgrade migrations and default rates. A high employment rate also should increase the probability of upgrade migrations.

Inflation rate: There exists a variety of price indices that measure inflation in an economy. We include the yearly percentage change in the seasonally adjusted Consumer Price Index (CPI) and Producer Price index (PPI). There is a common perception that inflation is bad for the economy. Thus, it is possible to have more default rates among firms. However, the real value of outstanding debt of a firm in terms of nominal dollars, could be reduced as a result of inflation. Therefore, it is possible to have lower rates of default among firms [see, Figlewski et al. (2012)].

Financial Markets

Interest rate: We include the short-term and the long-term Interest rates. The long-term interest rates are considered as a measure of the overall level of interest rate at longer maturity in an economy and the short-term interest rate is considered as a measure of the tightness of money market. Duffie et al. (2007) found a negative relation between firm's credit default and the short-term interest rate. The negative relation can be explained by the fact that central banks increase the short rates to cool down business expansion. We tend to have more optimistic markets in an economy with higher long-term interest rates.

Spread: We include the spread between the short and long-term interest rates.

Exchange rate: Firms with a high volume of international business are expected to be more sensitive to changes in exchange rates. If the exchange rate is high, the importing businesses are expected to be positively affected and exporting businesses are negatively affected.

Stock Markets

CAC 40 return: The general health of the corporate sector is indicated by a high performance of the stock market. Moreover, the level and volatility of firm's stock price have a direct effect on a firm's default risk exposure in a structural model. We include the return on the CAC 40 index. The probability of default and downgrade is negatively related to the stock market return.

CAC 40 volatility: We include the volatility index of CAC 40 as a measure of market risk. The probability of default is positively related to the volatility.

Composite Index

Composite leading indicator (CLI): There exists previous studies that showed a strong link between some composite indices and corporate default risk. Therefore, we choose the composite leading indicator (CLI) as a short-term measure of economic movement in qualitative term. The early signals of turning points in business cycles showing fluctuation of the economic activity around its long term potential level can be provided by the CLI.

These variables are collected from various sources including the St. Louis Federal Reserve Economic Data, the World Bank database and OECD database for France over the period of 2007-2014. These variables are shown in Table 2.15. The descriptive statistics of these macroeconomic variables are given in Appendix B.

Table 2.15: Description and Sources of Macroeconomic Variables

Category	**Variables**	**Description**	**Resources**
National accounts	GDP	Real GDP Growth Rate (annual %)	World Bank
	OUTG	Output Gap, % of Potential GDP (Actual GDP minus Potential)	St. Louis Federal Reserve Economic Data
	UNEM	Unemployment Rate (% of total labor force)	OECD database
	EMP	Employment Rate (% of total labor force)	OECD database
	IPROD	Production of Total Industry	OECD database
	CPI	Consumer Price Index	St. Louis Federal Reserve Economic Data
	PPI	Producer Price Index	St. Louis Federal Reserve Economic Data
Finantial markets	SINTS	Short-term Interest Rate	OECD database
	LINTS	Long-term Interest Rate	OECD database
	SPL-S	Spread between Short and Long-term Interest Rate	OECD database
	EXCH	Exchange Rate	OECD database
Stock markets	CAC40	Return of CAC 40 Index	Yahoo finance
	VCAC4	Volatility of CAC 40 Index	Yahoo finance
Composite index	CLI	Composite Leading Index	OECD database

The next step is to calculate the correlation between the systematic factor and these variables. However, the time dimension T of our data set is small, therefore, the estimated correlation between the systematic factor and these variables may not be accurate. Moreover, due to the non-linear pattern of the systematic factor and these variables the sample correlation may not be sufficient. Thus, we do not rely just on the correlation and consider the graphical analysis as well.

First, we calculate the correlation between these variables in Table 2.15 and the systematic factor which are reported in Table 2.16. Next, we compare the evolution of variables that have patterns similar to the systematic factor in Figure 2.9. We consider those variables that have a similar pattern as the systematic factor.

Table 2.16: Correlation Between the Macroeconomic Variables and the Systematic Factor

	Macroeconomic variables	Correlation with the systematic factor	P-Value
National accounts	GDP	-0.41	0.31
	OUTG	**-0.76*****	**0.02**
	UNEM	0.53	0.17
	EMP	-0.3	0.46
	IPROD	**-0.73*****	**0.03**
	CPI	0.47	0.22
	PPI	0.12	0.77
Financial markets	**SINTS**	**-0.72*****	**0.04**
	LINTS	-0.57	0.13
	SPL-S	0.52	0.18
	EXCH	-0.27	0.5
Stock markets	CAC40	0.22	0.59
	VCAC40	0.33	0.41
Composite index	CLI	-0.51	0.18

In terms of correlation between the systematic factor and the variables, we observe that, the output gap and the total industrial production from the national accounts group, have a high negative correlation, significant at 5% with the systematic factor. From the financial markets group, the systematic factor is inversely correlated with the short-term interest rate and this correlation is significant at 5%. These three variables have statistically significant correlation with the systematic factor.

In terms of graphical analysis, we plot each variable against the systematic factor and compare the pattern of each variable with the systematic factor. Among all the variables, the patterns of the output gap, the total industrial production, the short-term interest rate and the annual return on the CAC 40 are similar to the pattern of the systematic factor. Henceforth, we focus on the variables which have significant correlation with the systematic factor and the variables that graphically have the same pattern as the systematic factor.

Let us first take a look at the evolution of the output gap, total industrial production, from the national accounts group, the short-term interest rate from the financial markets and the annual return on CAC 40 from the stock markets group in France over the period

2007-2014. When we plot the variables that have a negative correlation with the systematic factor, we multiply the systematic factor by (-1) in order to have a better vision of the factor path. We keep the factor unchanged in the case of CAC 40 which has a positive correlation with the factor.

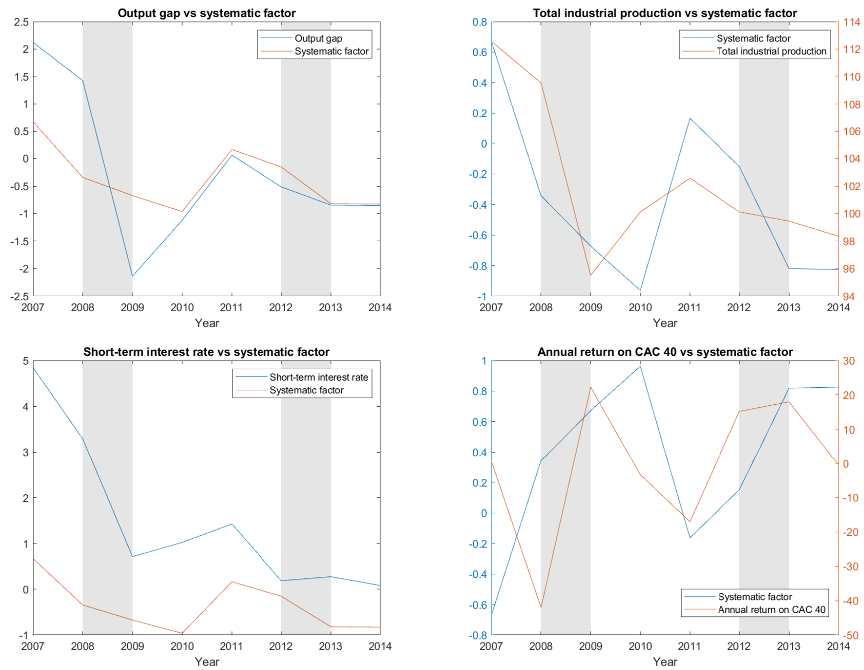


Figure 2.9: Correlation Between Macro-Variables and the Systematic Factor. This figure shows the evolution of the output gap % of Potential GDP, the total industrial production, the short-term interest rate, per cent, per annum (%), and the annual return on CAC 40 for France over the period 2007-2014 and the systematic factor. The shaded periods refer to OECD recessions in France.

It can be easily seen that the cyclical pattern of the negative factor is close to the output gap, the total industrial production and the short-term interest in France. Moreover, the systematic factor has a pattern similar to the annual return on the CAC 40. The output gap shows local downward trends during both recession periods in France, and since our factor is negatively correlated with the output gap, we observe local upward trends in the systematic factor during the recession periods. In 2011, the output gap has a peak and the systematic factor has a trough as well. We observe the same pattern for the total industrial production and the short-term interest rate as well. The CAC 40 index which is a benchmark French stock market index shows local upward trends during both recessions which coincide with upward trends in the systematic factor. Both series have a trough in 2011 as well.

In order to see how these variables are related to the systematic factor, we perform the linear regressions. These four aforementioned variables are regressed on the systematic factor. The results of the linear regressions are reported in Table 2.17.

Table 2.17: Linear Regression-Two step Estimation Method

Variables		Estimate	SE	t.Stat	Normal P-Value	R-Squared
OUTG	β_0	0.45	0.41	1.09	0.31	0.58
	β_1	$\mathbf{-1.87}^{***}$	0.64	-2.90	0.02	
IPROD	β_0	$\mathbf{105.03}^{***}$	1.84	57.07	0.00	0.53
	β_1	$\mathbf{-7.48}^{***}$	2.84	-2.63	0.03	
SINTS	β_0	$\mathbf{2.27}^{***}$	0.55	4.14	0.01	0.52
	β_1	$\mathbf{-2.17}^{***}$	0.85	-2.55	0.04	
CAC 40	β_0	-3.89	9.53	-0.4	0.69	0.05
	β_1	8.31	14.74	0.56	0.59	

The regression equation is as follows:

$$Y_t = \beta_0 + \beta_1 f_t + \epsilon_t,\ E(\epsilon_t) = 0,\ Var(\epsilon_t) = \sigma^2.$$

The slopes β_1 are statistically significant at 5% for the output gap, the total industrial production, and the short-term interest rate and the intercepts β_0 are statistically significant at 5% for the total industrial production and the short-term interest rate.

This regression analysis and the comparison of patterns of the variables displayed in Figure 2.9 suggest that the systematic factor obtained from the two-step efficient estimation approach is related to these variables. In addition to this analysis, we will perform the stress testing in the next Section which allows us to interpret the systematic factor in more detail.

Equivalently, the estimates of the micro- and macro parameters and the factor path can be derived from the joint optimization. In the next section we discuss the results of the joint optimization.

2.4.4 Joint Optimization Approach

In this Section, the factor values are treated as nuisance parameters.[5] The vector of parameters includes the micro-parameters θ and factor values f_t, and it is estimated by considering the solution in θ of the joint optimization problem given in equation (2.2.13). We use the estimates of the two-step estimation as initial values in the joint optimization approach. Table 2.18 displays the parameter estimates for the joint optimization method.

$\hat{c}_1 = -4.974^{***}$ (0.016)	$\hat{c}_2 = -1.851^{***}$ (0.001)	$\hat{c}_3 = 0$	$\hat{c}_4 = 0.671^{***}$ (0.0002)	$\hat{c}_5 = 1.587^{***}$ (0.001)	$\hat{c}_6 = 2.637^{***}$ (0.003)

$\hat{\delta}$		$\hat{\beta}$		$\hat{\sigma}$	
1	$\hat{\delta}_2 = -3.168^{***}$ (0.005)	1	$\hat{\beta}_2 = -0.016$ (0.026)	1	$\hat{\sigma}_2 = 0.977^{***}$ (0.003)
2	$\hat{\delta}_3 = -1.378^{***}$ (0.003)	2	$\hat{\beta}_3 = 0.379^{***}$ (0.010)	2	$\hat{\sigma}_3 = 0.888^{***}$ (0.001)
3	$\hat{\delta}_4 = 0$	3	$\hat{\beta}_4 = 1$	3	$\hat{\sigma}_4 = 1$
4	$\hat{\delta}_5 = 0.418^{***}$ (0.002)	4	$\hat{\beta}_5 = 0.360^{***}$ (0.007)	4	$\hat{\sigma}_5 = 1.158^{***}$ (0.002)
5	$\hat{\delta}_6 = 1.561^{***}$ (0.010)	5	$\hat{\beta}_6 = 0.881^{***}$ (0.022)	5	$\hat{\sigma}_6 = 1.275^{***}$ (0.001)
6	$\hat{\delta}_7 = 3.297^{***}$ (0.065)	6	$\hat{\beta}_7 = -1.851^{***}$ (0.245)	6	$\hat{\sigma}_7 = 3.896^{***}$ (0.068)

Table 2.18: Micro-Parameter Estimates-Joint Optimization Approach. Thresholds $\hat{c}_k$, $k = 1, ..., 6$, intercepts $\hat{\delta}_l$, factor sensitivities $\hat{\beta}_l$, and volatilities $\hat{\sigma}_l$, $l = 2, ..., 7$. Standard errors are given in parentheses. $*, p < 0.05$, $**$, $p < 0.01, ***$, $p < 0.001$

All the estimates are statistically significant except for $\hat{\beta}_2$. The upper panel shows the estimates of the threshold parameters c_k. As expected, they are increasing with regards to the rating class index. The estimates of the parameters in rows $l = 2, ..., 7$ of the transition matrix, which correspond to rating classes "D", "C", ..., "A+" are displayed in the lower

[5]This approach might create an incidental parameter problem since the number of nuisance parameters tends to infinity with T. However, we do not have this problem since the cross-sectional dimension is much larger than the time dimension.

panel. The intercepts δ_l are increasing with respect to the rating index. The estimates of factor sensitivities $\hat{\beta}_l$ are positive, except the factor sensitivity of the rating category "A+" and "D". Generally, the riskier rating categories have smaller idiosyncratic volatility σ_l.

The estimated factor values are reported in Table 2.19. The factor pattern obtained from the joint optimization is similar to the pattern of factor obtained by the two-step estimation. Figure 2.10 displays the pattern of estimated factor values. Recall that the factor estimates are standardized to obtain zero-mean and unit variance in the sample.

Table 2.19: Estimated Factor Values

Estimated factor values							
2007	2008	2009	2010	2011	2012	2013	2014
-1.5538	0.0769	0.5607	1.0736	-1.1349	-0.6761	0.8424	0.8112

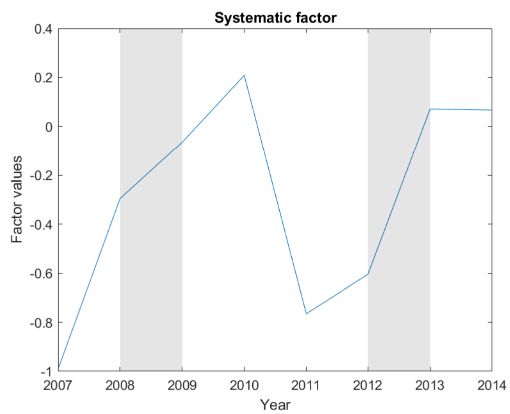


Figure 2.10: Systematic Factor-Joint Optimization Approach. The figure displays the pattern of the estimated factor values $\hat{f}_{n,T,t}$ for $t = 1, ..., 8$. The factor estimates are standardized to obtain zero-mean and unit variance in the sample. The shaded areas indicate the OECD recessions in France.

The factor series shows a local downward trend from 2010-2011 which is consistent with the local downward trend in the factor obtained from the two-step efficient estimation. The estimated macro-parameter $\hat{\rho}$, which is obtained from the maximum likelihood estimation of the autoregressive model (2.2.12) is reported in Table 2.20.

Table 2.20: Estimated Macro-Coefficients

	Estimate	SE	T.Stat	P-Value
μ	-0.2711	0.2152	-1.2596	0.2078
ρ	0.1203	0.4223	0.2659	0.7902
σ_η^2	0.1587	0.1843	0.8612	0.3891

The macro-parameter estimate $\hat{\rho}$ obtained from the joint optimization is positive and not statistically significant.

We perform a similar analysis to that in Section (2.4.3) to find the link between the factor pattern and the macroeconomic variables. The correlation between the systematic factor obtained by the joint optimization and the macro variables (Table 2.15) are shown in Table 2.21.

Table 2.21: Correlation Between Macroeconomic Variables and the Systematic Factor Obtained by the Joint Optimization Approach

	Macroeconomic variables	Correlation with the systematic factor	P-Value
National accounts	GDP	-0.36	0.36
	OUTG	$\mathbf{-0.69}^{***}$	**0.05**
	UNEM	0.47	0.23
	EMP	-0.25	0.54
	IPROD	$\mathbf{-0.65}^{*}$	**0.08**
	CPI	0.38	0.35
	PPI	0.01	0.96
Financial markets	SINTS	-0.61	0.1
	LINTS	-0.5	0.2
	SPL-S	0.44	0.27
	EXCH	-0.17	0.67
Stock markets	CAC40	0.19	0.64
	VCAC40	0.27	0.51
Composite index	CLI	-0.5	0.2

The systematic factor obtained by the joint optimization has a negative significant correlation with the output gap (real GDP actual minus potential) and the total industrial production in France. Similar to our regression analysis in Section (2.4.3), the output gap and the total industrial production are regressed on the systematic factor. The results of the linear regressions are reported in Table 2.22.

Table 2.22: Linear Regression-Joint Optimization Approach

Variables		Estimate	SE	t.Stat	Normal P-Value	R-Squared
OUTG	β_0	-0.88	0.46	-1.88	0.10	0.48
	β_1	$\mathbf{2.20}^{***}$	0.91	-2.39	0.05	
IPROD	β_0	$\mathbf{99.74}^{***}$	2.07	48.01	0.00	0.42
	β_1	$\mathbf{-8.54}^{*}$	4.06	-2.09	0.08	

The slopes β_1 are statistically significant at 5% for the output gap and at 10% for the total industrial production. The intercept β_0 for the total industrial production is statistically significant at 5%. The regressions suggests that the systematic factor obtained by the joint optimization approach might be related to these two variables. By comparing the pattern of the factor estimated by the two approaches, and given the parameter estimates from each approach, we conclude that, we have found links between the systematic factor obtained by both estimation methods and some national accounts variables, like the output gap and the total industry production and the financial markets variable, which is the short-term interest rate, in France. In the next Section, we compute the fitted transition probabilities from the two-step and the joint optimization parameter estimates and compare them with the observed migration probabilities.

2.5 Fitted Credit Migration Matrices

To assess the fit of the model, we compare the observed and estimated transition matrices. The Euclidean distance or Euclidean metric is one way to compare two matrices. This mathematical method is based on the straight-line distance between two points in the Euclidean space. This method computes the average absolute difference of the corresponding elements of two matrices. Moreover, one can measure the average root-mean-square difference between multiple corresponding elements of the matrices [see, e.g. Israel et al. (2001), Bangia et al.

(2002)]. However, the Euclidean metric provides a relative measure of distance between two matrices and it is difficult to judge whether this distance is large or small. Moreover, it is not easy to interpret the economic meaning of the distance. Another method which allows us to compare two transition matrices is a mobility index [see, Geweke et al. (1986)]. Jafry and Schuermann (2004) extend that method and introduce a singular value metric (SVD-mobility index) which is a function that compares two migration matrices with respect to their ability to generate migration events.

In order to assess and compare which estimation procedure delivers the best fitted migration probabilities, we proceed in two steps. First, we compute the estimated transition probabilities from both approaches and compare the stability rates, downgrade rates and upgrade rates of each rating category implied by each method during the study period with the empirical stability rates, downgrade and upgrade rates for each rating class. Next, we calculate the distance between the observed and fitted probabilities to see how close the fitted probabilities are to the observed migration frequencies obtained by each estimation approach. Second, we compare the transition matrices implied by each estimation methods by using the SVD-mobility index introduced by Jafry and Schuermann (2004). The estimated transition matrices obtained by the two-step efficient estimation method and the joint optimization approach for the whole sample period are reported in Appendix C.1 and C.2 respectively.

2.5.1 Comparison using the Stability, Downgrade and Upgrade Rates

In-Sample Estimates of Migration Probabilities-Two-step Estimation Approach

First, we compute the implied probabilities of staying in the same rating class known as the implied stability rates. The evolution of the implied stability rates and the empirical stability rates for each rating class are presented in Figure 2.11.

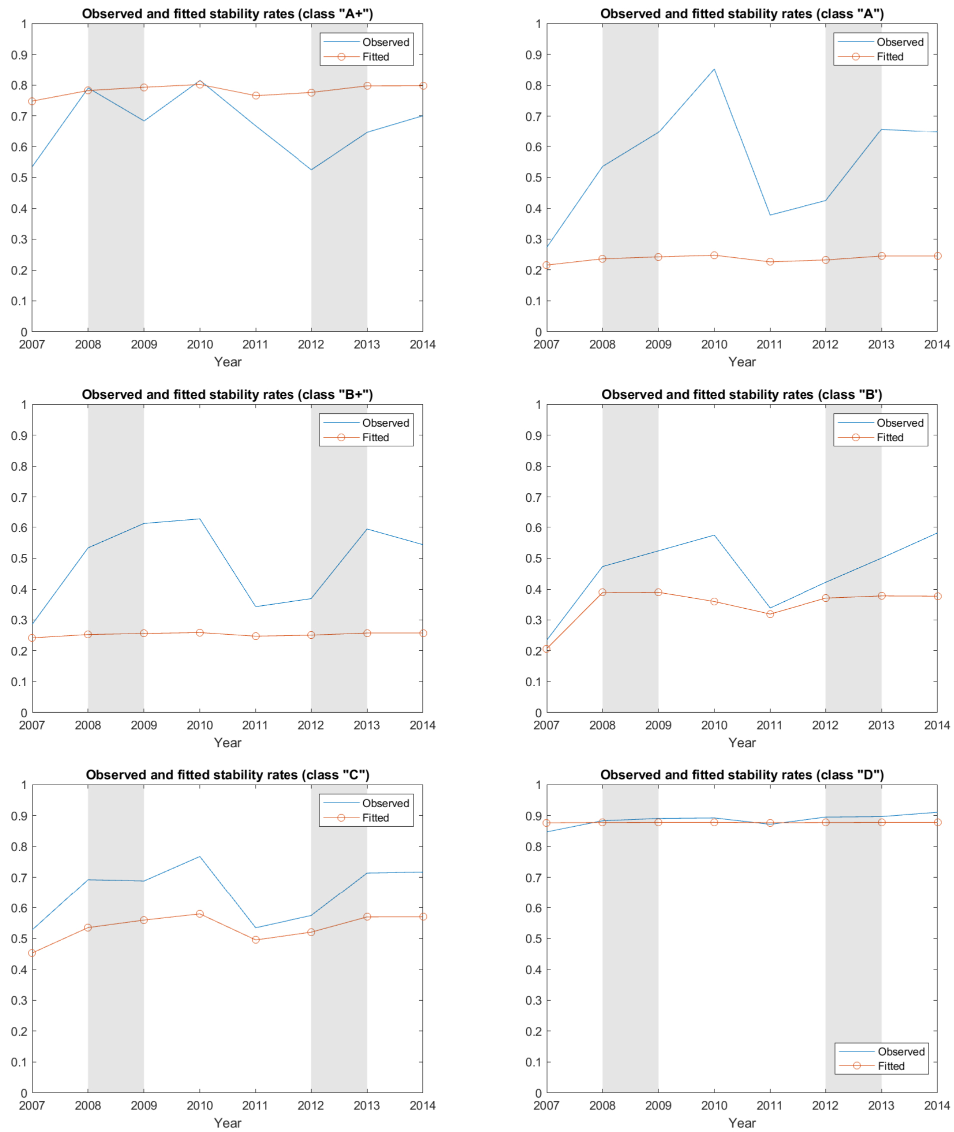


Figure 2.11: Observed and Fitted Stability Rates. This figure shows the time series of observed probabilities of staying in the same class (solid line) and fitted probabilities of staying in the same class obtained by the two-step estimation approach (circles). The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

The pattern of the fitted stability rates captures the pattern of observed stability rates especially in the rating classes “B”, “C”, and “D”. Figure 2.12 displays the observed and fitted downgrade probabilities for each rating class.

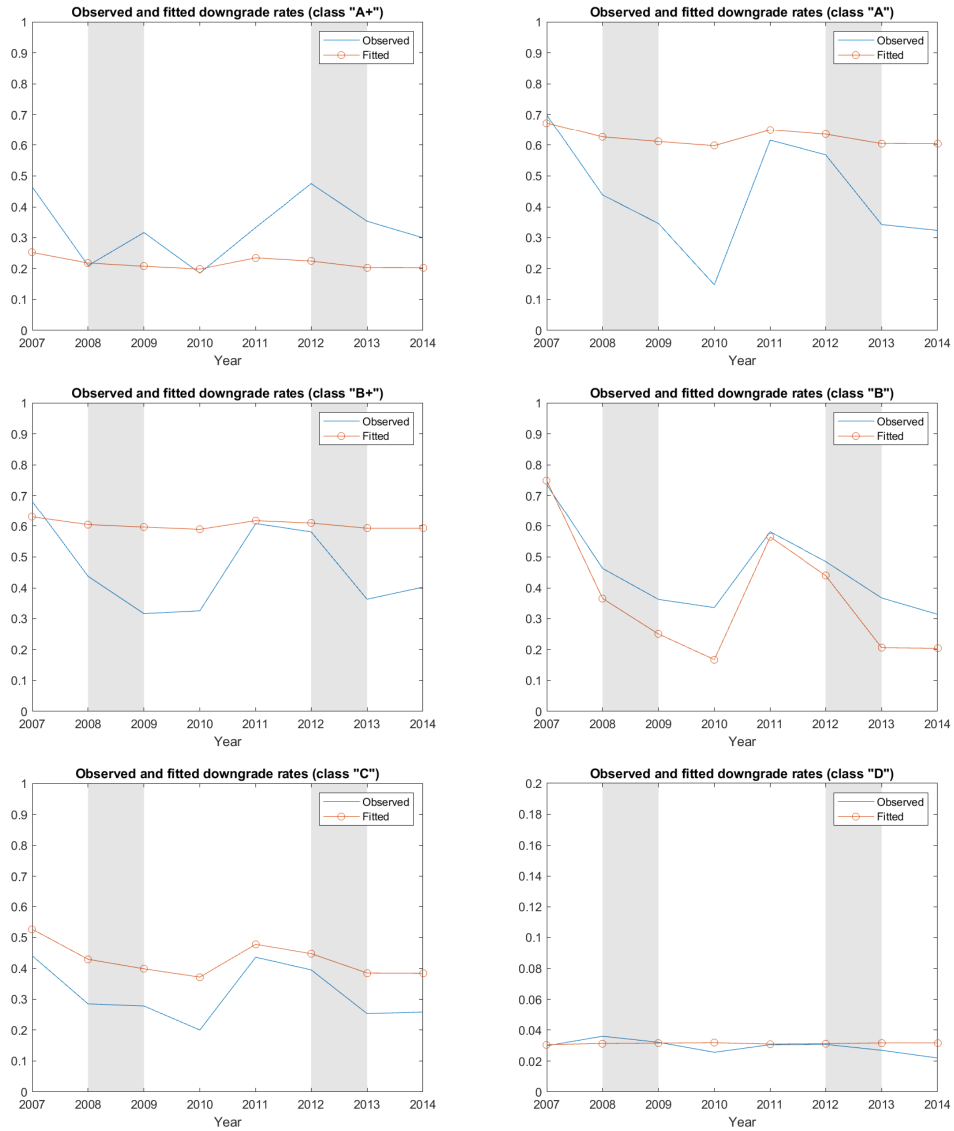


Figure 2.12: Observed and Fitted Downgrade Rates. This figure shows the time series of observed probabilities of downgrade for each rating class (solid line) and fitted probabilities of downgrade for each rating class obtained by the two-step estimation approach (circle). The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

The pattern of the fitted downgrade probabilities in all rating classes captures the pattern of observed downgrade probabilities which confirms that our model is successful in capturing the downgrade risk. Figure 2.13 demonstrates the observed and fitted upgrade probabilities for each rating class.

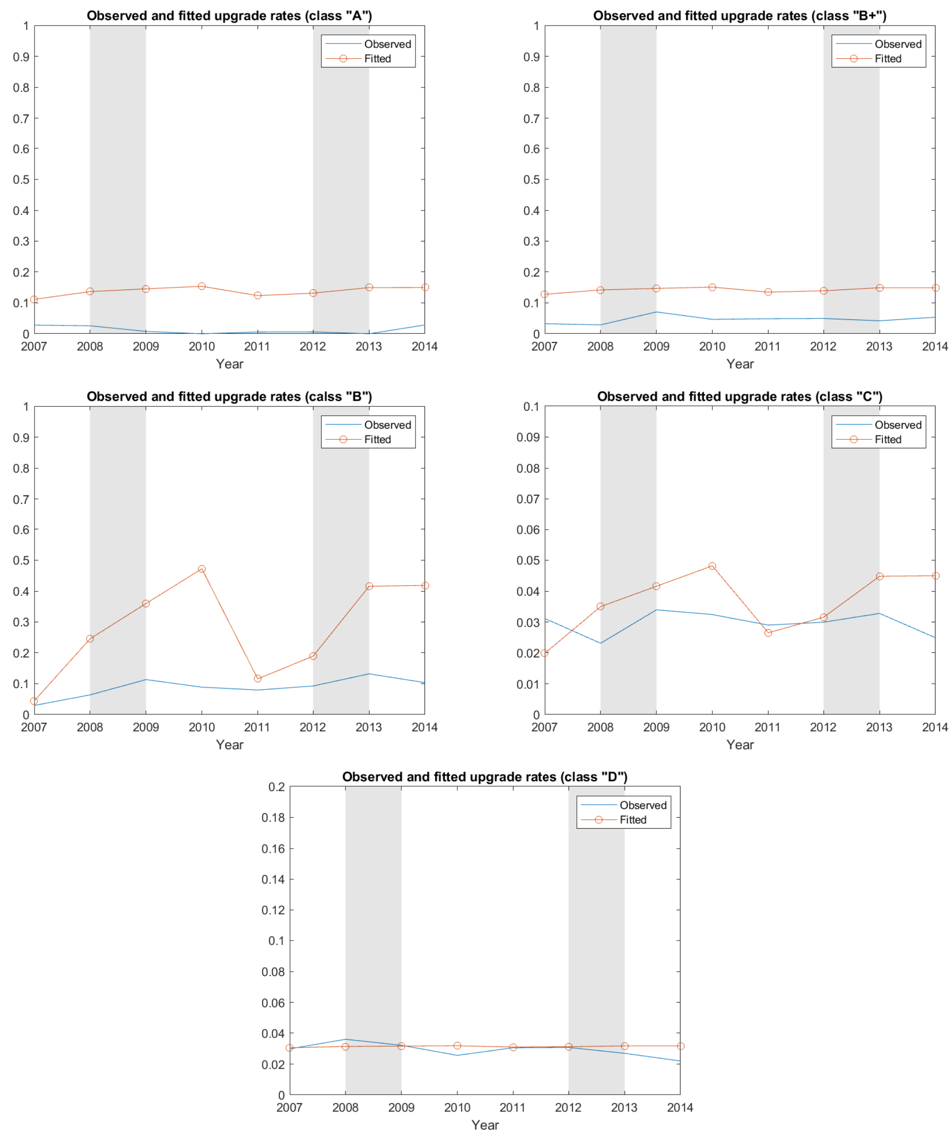


Figure 2.13: Observed and Fitted Upgrade Rates. This figure shows the time series of observed probabilities of upgrade for each rating class (solid line) and the fitted probabilities of upgrade for each rating class obtained by the two-step estimation approach (circle). The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

Moreover, some actual and fitted migration probabilities from rating class “B”, $l = 4$, to rating categories “A”, “B+”, “C”, and “D”, $k = 6, 5, 3, 2$ are displayed in Figure 2.14. The peaks and troughs in the observed and fitted time series of migration probabilities for rating class “B” to the ratings “A”, “B+”, “C”, and “D” coincide in all panels of Figure 2.14.

Figure 2.14: Time series of Observed and Fitted Migration Probabilities from Rating Classes “B” to Rating Categories “A”, “B+”, “C”, and “D”, $k = 6, 5, 3, 2$. The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

Figures 2.11-2.14 indicate that the fitted probabilities capture the trends in the observed migration probabilities although, their values may differ. Next, we calculate the average vertical distance between the observed and fitted migration probabilities as follows:

$$Dis = \frac{1}{T}\sum_{t=1}^{T}(p_{lk,t} - \hat{p}_{lk,t})^2.$$

Table 2.23 shows the average distance between the observed and fitted stability, downgrade and upgrade rates.

Table 2.23: Average Distance Between Observed and Fitted Migration Probabilities

	A+	A	B+	B	C	D
Stability rates	0.02	0.12	0.07	0.01	0.01	0.0003
Downgrade rates	0.02	0.05	0.03	0.01	0.01	2.26E-05
Upgrade rates	0	0.01	0.009	0.05	0.0001	0.0002

The best fitted downgrade and upgrade rates are for rating classes "D" and "C", respectively. The one-factor model is successful in reproducing the general pattern of the expected risk in each rating class. The relatively higher discrepancies for some rating classes, may be due to the fact that some of the assumptions of the model may not hold. For example, the assumption of a homogeneous population of firms is a strong assumption. Another reason could be the choice of the cumulative distribution function Φ, which can be different from the Gaussian one, and feature different tail or skewness behaviors. Moreover, the time dimension T of the panel is short, which affects the accuracy of the estimators. Let us take a look at the estimated migration probabilities in the joint optimization approach. In the next section, we compare the fitted and observed migration probabilities obtained from the joint optimization.

In-Sample Estimates of Migration Probabilities-Joint Optimization Approach

In order to compare the fit provided by the joint optimization with the two-step efficient estimation approach, we compute the fitted stability rates, downgrade, upgrade and default rates for each rating class and compare them with the empirical probabilities. Figure 2.15 displays the observed and fitted stability rates for each rating class from the joint optimization approach.

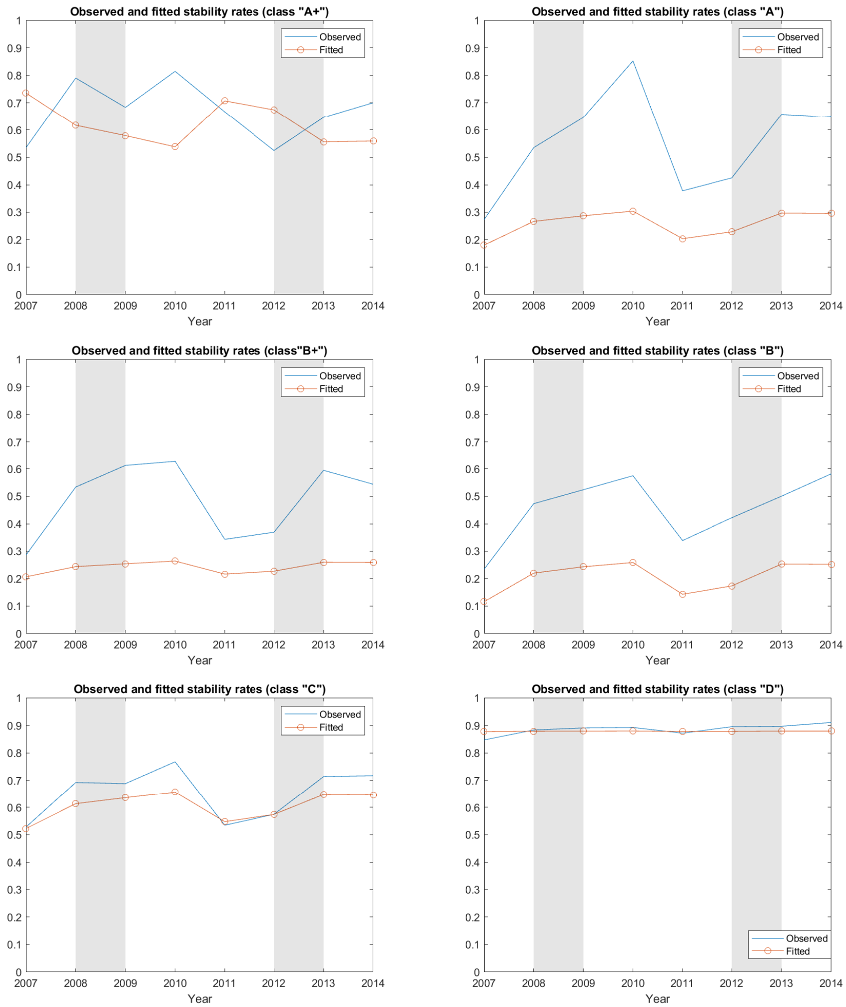


Figure 2.15: Observed and Fitted Stability Rates. This figure shows the time series of observed probabilities of staying in the same class (solid line) and fitted probabilities of staying in the same class obtained by the joint optimization approach (circles). The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

By visual inspection, the pattern of fitted stability rates from the joint optimization approach is close to the pattern of the fitted stability rates obtained by two-step estimation approach. Figures 2.16 and 2.17 display the fitted downgrade and upgrade probabilities respectively.

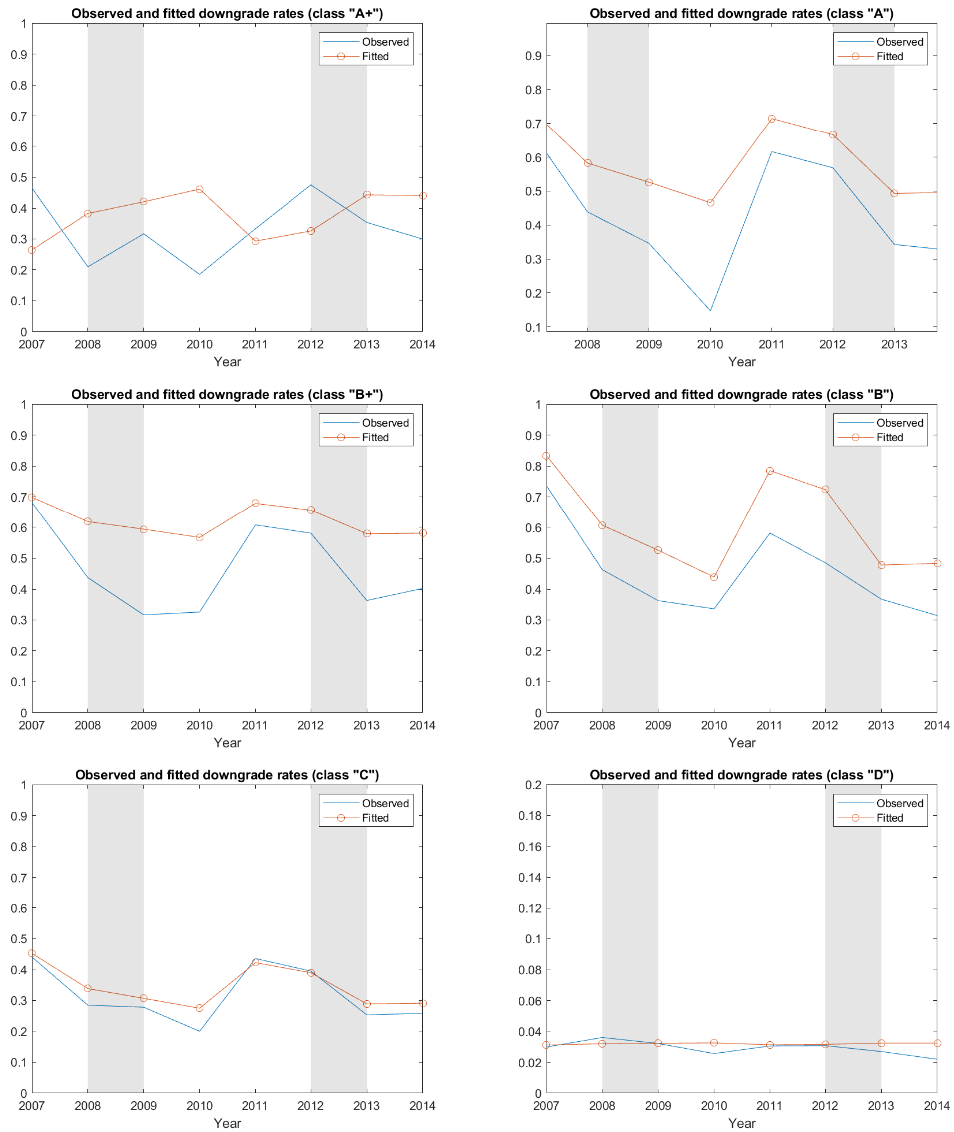


Figure 2.16: Observed and Fitted Downgrade Probabilities. This figure shows the time series of observed probabilities of downgrade for each rating class (solid line) and fitted probabilities of downgrade obtained by the joint optimization approach for each rating class (circle). The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

The patterns of the fitted downgrade probabilities from the joint optimization approach are close to the patterns of observed downgrade probabilities in all rating classes except for rating categories “B+”.

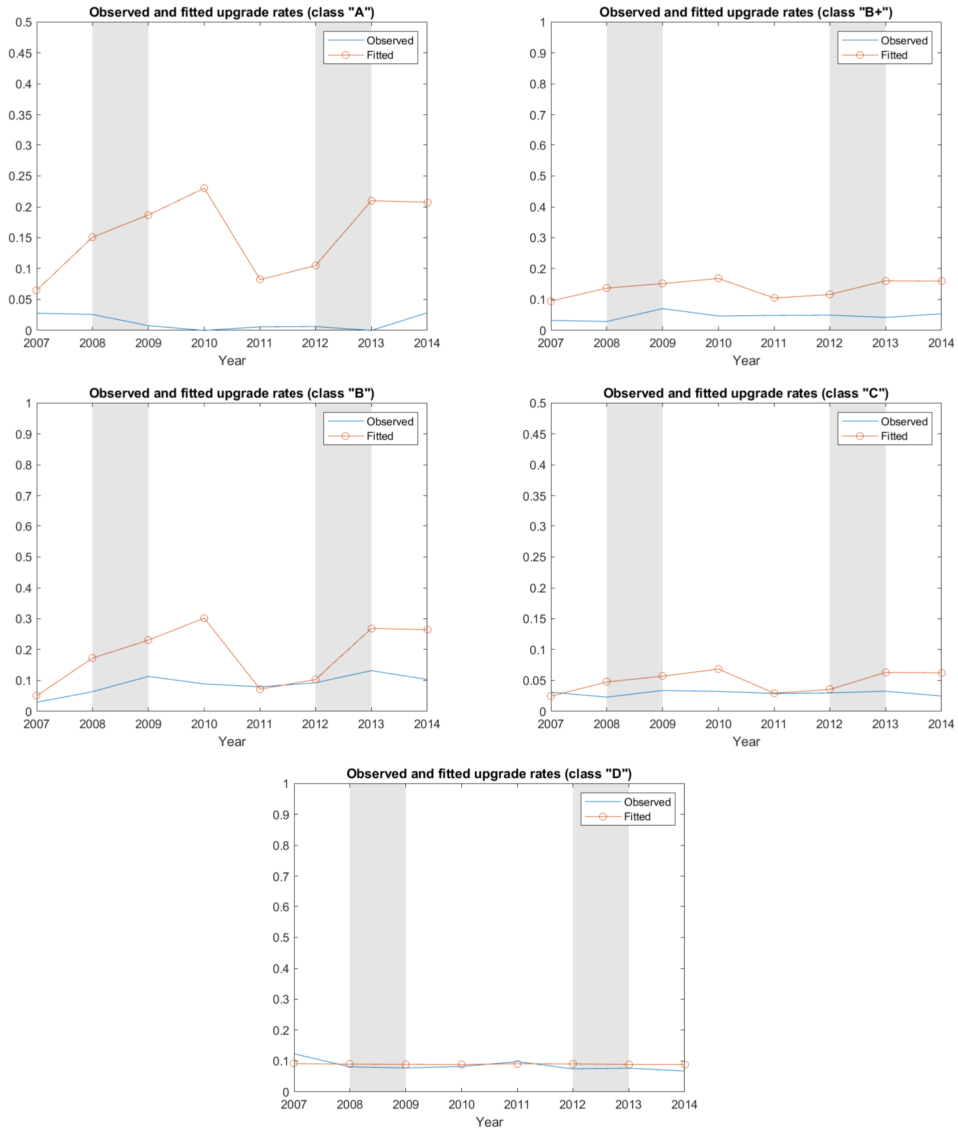


Figure 2.17: Observed and Fitted Upgrade Probabilities. This figure shows the time series of observed probabilities of upgrade for each rating class (solid line) and the fitted probabilities of upgrade obtained by the joint optimization approach for each rating class (circle). The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

The patterns of the fitted upgrade probabilities from the joint optimization approach are close to the patterns of observed upgrade probabilities in all rating classes except for rating categories "A+".

Next, we plot the fitted migration probabilities from rating class "B", $l = 4$, to rating categories "A", "B+", "C", and "D", $k = 6, 5, 3, 2$, in Figure 2.18. The probabilities are close to the observed frequencies.

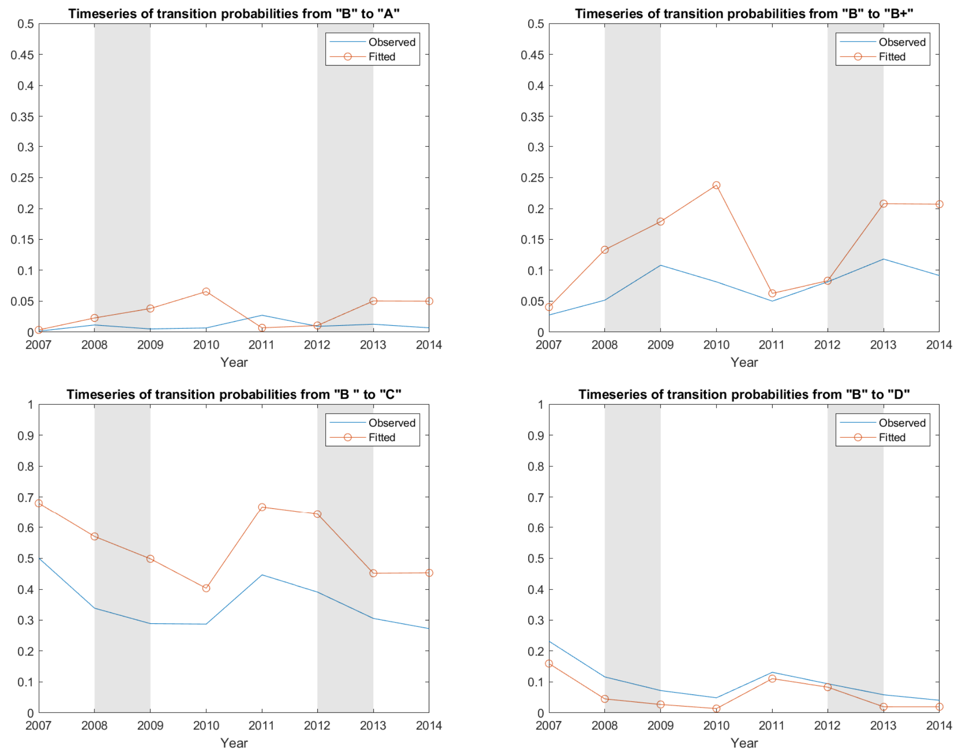


Figure 2.18: Time series of Observed and Fitted Migration Probabilities from Rating Classes "B" to Rating Categories "A", "B+", "C", and "D", $k = 6, 5, 3, 2$ obtained by the joint optimization. The migration probabilities are given in percentages on the vertical axes. The shaded areas indicate the OECD recessions in France.

Table 2.24 shows the average distance between the observed and fitted stability, downgrade and upgrade rates obtained by the joint optimization approach.

Table 2.24: Average Distance Between Observed and Fitted Migration Probabilities

	A+	A	B+	B	C	D
Stability rates	0.03	0.11	0.07	0.06	0.004	0.0003
Downgrade rates	0.03	0.04	0.03	0.02	0.001	2.4021e-05
Upgrade rates	0	0.01	0.009	0.01	6.3253e-04	0.0002

The average distance between the observed and fitted stability rates, downgrade rates and upgrade rates obtained by the joint optimization is higher than from the two-step approach for rating classes "A+" and "B" smaller for rating classes "A" and "C" equal for "B+" and "D". The next Section compares these two estimation methods using the total distance between the fitted and the observed migration probabilities.

Comparison of estimation methods

We compute the goodness of fit measure for the two-step and the joint optimization approaches in each year as follows:

$$\chi_t^2 = \sum_{k=1}^{K}\sum_{l=1}^{K}(\hat{p}_{lk,t} - p_{lk,t}(\theta; f_t))^2,$$

In order to avoid the problem of null observed probabilities, the chi-square measure is not weighted by the inverse of the frequencies. Therefore, these measures do not asymptotically follow a χ^2 distribution and cannot be used for testing the distance [Feng et al. (2008)]. The χ^2-goodness of fit measures the total distance between the fitted and observed migration probabilities for the two-step estimation and the joint optimization approaches are reported in Table 2.25 and Table 2.26 respectively.

Table 2.25: Goodness of Fit by the Two-Step Efficient Estimation Approach

Year	2007	2008	2009	2010	2011	2012	2013	2014
χ^2	0.2226	0.3663	0.5129	0.8641	0.1193	0.2337	0.5322	0.5145

Table 2.26: Goodness of Fit by the Joint Optimization Approach

Year	2007	2008	2009	2010	2011	2012	2013	2014
χ^2	0.1608	0.4307	0.5151	0.8328	0.1682	0.2808	0.4867	0.5291

The χ^2-goodness of fit measures are larger from the joint optimization approach in all years except for years 2007, 2010 and 2013. Overall, the two-step approach provides the best fit, since the goodness of fit measures are smaller compared to the joint optimization approach.

2.5.2 Comparison using the SVD-Mobility Index

Jafry and Schuermann (2004) introduce a metric called the SVD-mobility index to compare two transition matrices on their ability to produces migration events. The SVD-mobility index is the most commonly used method of comparison for migration matrices in the literature. The mobility index provides a scalar value which captures the overall dynamic features of given transition matrices. This is the advantage of the mobility index as, instead of comparing the migration probabilities in each year, we can compare the overall dynamics of the matrices. We use Jafry and Schuermann (2004) SVD-mobility index to compare the migration matrices implied by the two-step estimation and the joint optimization methods with the observed migration matrices in the data.

The SVD-mobility index $M_{SVD}(\mathbb{P})$ is defined as a function of transition matrix $\mathbb{P}$. The main feature in a transition matrix $\mathbb{P}$ is the amount of migration or "mobility" imposed on the state vector from one period to the next. We subtract the identity matrix $\mathbb{I}$ from the original matrix $\mathbb{P}$, to get the mobility matrix $\tilde{\mathbb{P}}$. The mobility matrix includes only the dynamic part of the original matrix $\mathbb{P}$ and reflects the "magnitude" of $\mathbb{P}$ in terms of the the implied mobility since the main diagonal contains the negative values of the sum of the row elements. Jafry and Schuermann (2004) find that the average of all the singular values of mobility matrix $\mathbb{P}$ captures the general characteristics of that matrix. The SVD-mobility index for a d-dimensional matrix $\mathbb{P}$ is given by:

$$M_{SVD}(\mathbb{P}) = \frac{\sum_{i=1}^{d} \sqrt{\lambda_i(\tilde{\mathbb{P}}'\tilde{\mathbb{P}})}}{d},$$

where $\tilde{\mathbb{P}} = \mathbb{P} - \mathbb{I}$, and $\mathbb{P}'$ is its transpose. $\lambda_i(\tilde{\mathbb{P}}'\tilde{\mathbb{P}})$ is the i^{th} eigenvalue of $\tilde{\mathbb{P}}'\tilde{\mathbb{P}}$, sorted in decreasing order, i.e. $\lambda_1(\tilde{\mathbb{P}}'\tilde{\mathbb{P}}) > .... > \lambda_d(\tilde{\mathbb{P}}'\tilde{\mathbb{P}})$.

Suppose that the values of all diagonal elements of matrix $\mathbb{P} = \mathbb{P} - \mathbb{I}$ are $(1-p)$ and all off-diagonal elements are equal to $(p/d-1)$, where p represents the probability that a given state will undergo a migration (to any of the others). Then, Jafry and Schuermann (2004) show that the M_{SVD} metric applied to $\mathbb{P}$ yields the following exact result:

$$M_{SVD}(\mathbb{P}) = p.$$

This result indicates that the average migration probability p is numerically identical to the average singular value metric. For instance, if we get the value 0.1 of the singular value metric for a given matrix, say, then we can conclude that the matrix has an effective average probability of migration of 0.1. Thus, the difference between two migration matrices could be measured by M_{SVD} as follows:

$$D_{SVD}(\mathbb{P}_1, \mathbb{P}_2) = M_{SVD}(\mathbb{P}_1) - M_{SVD}(\mathbb{P}_2).$$

A directional deviation between two matrices in terms of the mobility or approximate average probability of migration can be measured by this difference. To compare the mobility "size" of the migration matrices, we use the M_{SVD} index for the period 2007-2015 for the two-step estimation and the joint optimization methods. We apply the SVD mobility index to each matrix calculated according to the two-step and the joint optimization approaches. The results are shown in the Tables 2.27 and 2.28.

Table 2.27: M_{SVD}-Mobility Index by the Two-Step Estimation Approach

Year	2007	2008	2009	2010	2011	2012	2013	2014
Two-steps	0.2577	0.2257	0.2214	0.2212	0.2400	0.2302	0.2209	0.2209
Observed data	0.2465	0.1581	0.1435	0.1145	0.2163	0.2063	0.1478	0.1437
D_{SVD}	0.0111	0.0676	0.0779	0.1067	0.0236	0.0239	0.0731	0.0772

Table 2.28: M_{SVD}-Mobility Index by the Joint Optimization Approach

Year	2007	2008	2009	2010	2011	2012	2013	2014
Joint optimization	0.2675	0.2412	0.2348	0.2301	0.2608	0.2532	0.2319	0.2322
Observed data	0.2465	0.1581	0.1435	0.1145	0.2163	0.2063	0.1478	0.1437
D_{SVD}	0.0210	0.0830	0.0913	0.1155	0.0444	0.0469	0.0841	0.0885

We can see that the metrics estimated from the joint optimization are slightly "larger" than the metrics for the two-step estimation method. One explanation for the relatively higher values in the joint optimization method could be the fact that the joint optimization method has a higher number of parameters to estimate, as compared to the two-step estimation which leads to less efficient estimators. Figure 2.19 shows the annually computed $M_{SVD}(P)$-mobility indexes of each estimation method and the observed data.

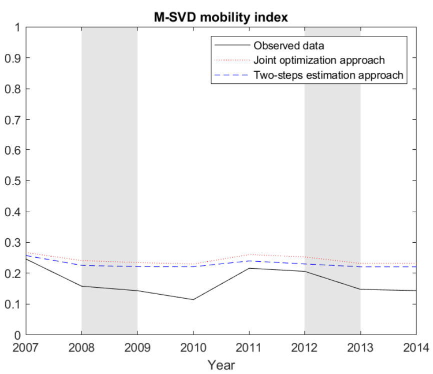


Figure 2.19: The M_{SVD} Mobility Index. This Figure exhibits the annual mobility of the migration matrices using the M_{SVD} for the period of 2007-2014. The two-step estimation approach: dotted line, the joint optimization method: dashed line and the observed data: solid line. The shaded areas indicate the OECD recessions in France.

According to the M_{SVD} both estimation methods over-estimate the empirical transition matrices. The $D_{SVD}(\mathbb{P}_1, \mathbb{P}_2)$ between each estimation approach and the M_{SVD} index of the empirical data are displayed in Figure 2.20, which shows that the differences of these two estimations methods are very small across time. The differences seems to increase during the recession of 2008 and prior to the recession of 2012.

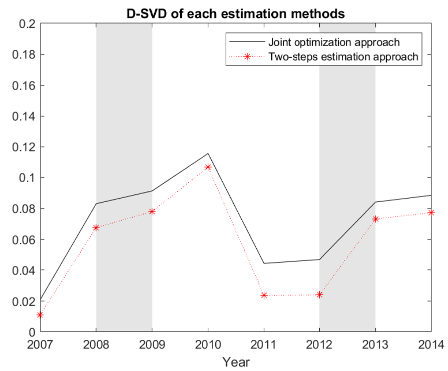


Figure 2.20: The D_{SVD} of Each Estimation Methods. This Figure exhibits the differences between M_{SVD} mobility index of the two-step estimation approach and M_{SVD} mobility index of the empirical data: solid line and the deference between M_{SVD} mobility index of the joint optimization approach and M_{SVD} mobility index of the empirical data: dotted line. The shaded areas indicate the OECD recessions in France.

In this section we performed the empirical comparison between the the two-step estimation method and the joint optimization approach. First, we compared the the average vertical distance between the observed and fitted stability, downgrade and upgrade rates in these two methods. Based on these comparison, The average distance of the stability, downgrade and upgrade rates between the observed and the fitted migration probabilities obtained from the two-step estimation method is less than the average distance obtained from the joint optimization approach. Next, we computed the goodness of fit for each year for these two methods. We found that the total distance between the fitted and observed migration probabilities for the two-step estimation approach is smaller in the two-step estimation method compare to the joint optimization approach. The last comparison was based on the SVD-mobility index which captures the overall dynamic features of given transition matrices. The result of the SVD-mobility index reveals that the metrics estimated from the joint optimization are slightly "larger" than the metrics for the two-step estimation method. Overall, we can conclude that the two-step and the joint optimization methods fit well the frequencies of transition matrices as the M_{SVD} measures are close to the empirical M_{SVD} values. However, the we conclude that the two-step estimation method provides better fit of the migration probabilities compare to the joint optimization approach. Moreover, the two-step method computation is less time consuming than the joint optimization method and the two-step estimation estimators have less standard deviation since the number of the parameter to be estimated are less in the two-step estimation method.

In the next step, we perform the stress testing analysis in the framework of the factor stochastic migration model. We apply the stress testing analysis to the stochastic migration model estimated by the two-step approach.

2.6 Stress Testing with the Stochastic Migration Model

As stated earlier in the text, estimations of probabilities of default (PD) and migration rates under hypothetical or historical stress scenarios are required by Basel II. Stress testing is an important tool to assess financial stability of banks and financial institutions. Conducting regular regulatory stress testing exercises under the supervision of the European Banking Authority (EBA) is mandatory for financial institutions in Europe[6]. The aim is to determine the behavior of bank portfolios under stress conditions. In particular, the impact of extreme events or changes in market conditions on the capital adequacy of banks, based on the scenarios of macroeconomic and financial shocks can be assessed by the stress testing analysis. Generally, credit risk stress tests concern risk parameters such as the probability of default at a one-year risk horizon (PD), migration probability, loss given default (LGD) or exposure loss at Default (EAD). These risk measures form the main building blocks of the Basel II regulatory capital supervision for credit risk. In our particular framework, the stress test exercise focuses on the probability of migration from one credit rating grade to another, and especially, the probability of migrating to default.

We study the changes in the migration rates in our portfolio, if the economy experiences a moderate downturn or upturn. For this purpose a one-time shock is added to the estimated systematic risk factor which affects the migration rates. The following hypothetical scenarios are considered:

[6] OSFI is the supervisory authority (Office of the Supervisor of Financial Institution) in Canada

- Scenario 1: A positive one-time shock to the systematic factor f_t in which the factor increases by one standard deviation of factor σ_f.

- Scenario 2: A negative one-time shock to the systematic factor f_t in which the factor decreases by one standard deviation of factor σ_f.

These scenarios are designed to resemble an improvement and deterioration of the economic situation. Furthermore, we consider two economic principles, "mutatis mutandis", ie "allowing other things to change accordingly" and "ceteris paribus", ie "all other things being equal" or "holding other factors constant" in our stress testing analysis. In the first case, i.e. "mutatis mutandis", the one-time positive and negative shocks are added to the systematic factor in years 2008 and 2010 separately. These shocks have an effect on the systematic factor in those years as well the factor values over the following years. In other words, we assume that the factor in years following the shocks are returns to the estimated trajectory. In the other case, i.e. "ceteris paribus", we impose the two shocks to the systematic factor in years 2008 and 2010, then, assume that over the next years after the shocks, the factor returns to a flat baseline trajectory. We examine the effects of shocks to determine how quickly each shock and its effects dissipate.

We consider these two above assumptions and consider the positive and negative shocks to the systematic factor in years 2008 and 2010 separately. The stressed migrations under both assumptions and shock scenarios computed from the two-step estimation for the periods affected by shocks are reported in Appendices D.

2.6.1 Shocks to the Systematic Factor- Assuming "mutatis mutandis"

We add a positive one-time and a negative one-time shock equal to one standard deviation of factor, σ_f, to the systematic factor in years 2008 and 2010. Figures 2.21 and 2.22 show the paths of the systematic factor and of the stressed factor under both stress scenarios in years 2008 and 2010 respectively.

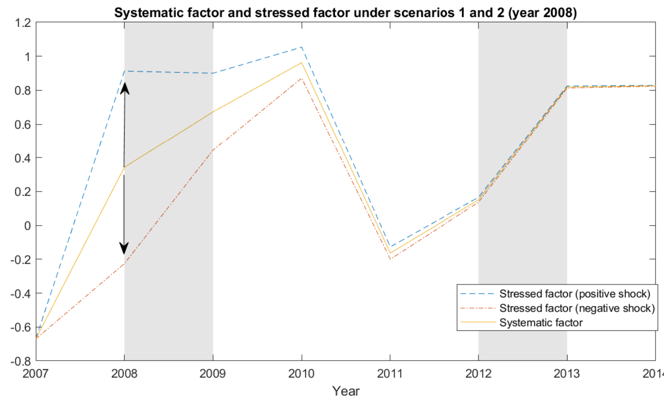


Figure 2.21: Systematic Factor and Stresses Factor Under Scenarios 1 and 2 (Year 2008). This Figure exhibits the systematic factor and the stressed factor under both scenarios, shocks are applied in 2008. The shaded areas indicate the OECD recessions in France.

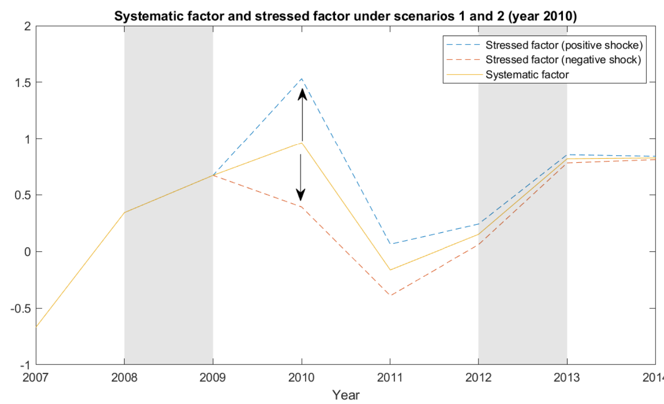


Figure 2.22: Systematic Factor and Stresses Factor Under Scenarios 1 and 2 (Year 2010). This Figure exhibits the systematic factor and the stressed factor under both scenarios, shocks are applied in 2010. The shaded areas indicate the OECD recessions in France.

When we impose the two shocks on the systematic factor in 2008, the effects of shocks are significant over the next two periods. Afterwards, the effects dissipate gradually. We observe the same pattern when we impose the shocks on the systematic factor in 2010. The shocks seem to dissipate faster when applied in 2010, i.e. the economy after a recession.

Let us compare the total default rates for all rating classes. We compute the stressed migration matrices from 2008 to 2014 and compare the estimated total default rates by all rating categories and stressed total default rates by all rating classes in years 2008 to 2014. Next we compare the stability rates, downgrade and upgrade probabilities. Figure 2.23 displays the estimated and stressed total default rates when we impose positive or negative shocks to the systematic factor in 2008.

Once we impose the positive shock on the factor, the total default rates of all rating classes decrease over the years 2008 to 2013. We observe 0.61% reduction in the total default rates in 2008 and 0.2% in 2010. The effect of positive shock is stronger in 2008 than in 2009. Afterwards, the total default rates are decreasing until 2011 and the effects of shock disappear. Then, the shocked default rates go back to the estimated levels of default rates.

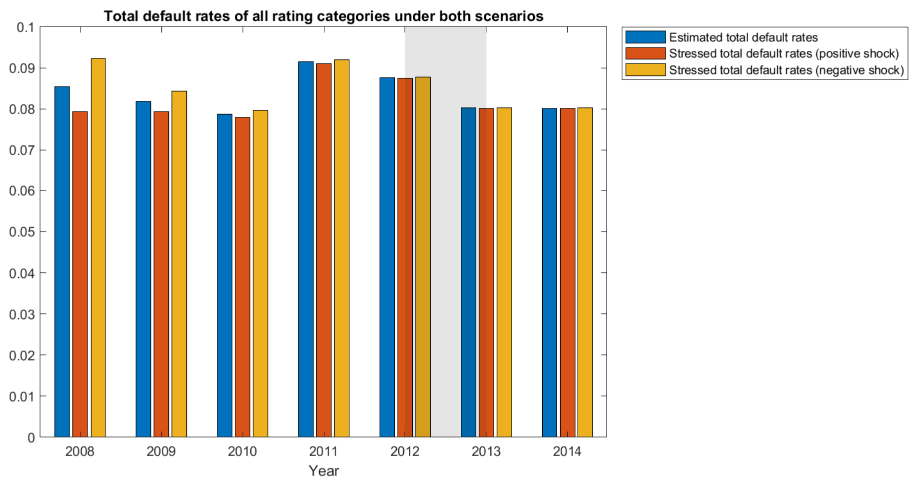


Figure 2.23: Total Default Rates of All Rating Classes Under Both Scenarios (2008). This Figure shows the estimated total default rates of all rating categories (blue) for years 2008 to 2014, the stressed total default rates after positive shock (red) and the stressed total default rates after negative shock (yellow). The shaded areas indicate the OECD recessions in France.

As a consequence of the one-time negative shock to the systematic factor in year 2008, the total default rates of all rating categories increase over the period 2008 to 2013. We observe an increase of 0.68% in 2008 and of 0.25% in 2009. The effect of the negative shock is stronger in 2008 than in 2009. Afterwards, the total default rates keep increasing until

2011. Then, the effects of shock disappear and the shocked default rates go back to the estimated levels of default rates. Next, let us compare the downgrade rates. Table 2.29 shows the changes in downgrade rates for each rating category over the years 2008 to 2014 under both a positive and a negative shocks to the systematic factor in 2008.

The downgrade rates for all rating classes decrease when we impose the positive shock on the factor, and increase with the one-time negative shock to the systematic factor in 2008 except for firms rated "D". For instance, the downgrade rates of firms rated "A+" have decreased by 1.8% at the time we imposed the positive shock. They increased by almost 2% when we imposed the negative shock on the factor. The effect of the positive and negative shocks are much stronger in the first two years after the shocks as compared to other years.

Table 2.29: The Changes in Downgrade Rates

Rating class	changes in the downgrade rates						
	One-time positive shock to the factor in year 2008						
	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
A+	-0.0180	-0.0071	-0.0028	-0.0012	-0.0005	-0.0002	-0.0001
A	-0.0264	-0.0106	-0.0043	-0.0016	-0.0007	-0.0003	-0.0001
B+	-0.0143	-0.0057	-0.0023	-0.0009	-0.0004	-0.0001	-0.0001
B	-0.1848	-0.0666	0.0219	-0.0144	-0.0057	-0.0017	-0.0007
C	-0.0531	-0.0211	-0.0083	-0.0035	-0.0014	-0.0005	-0.0002
D	0.00044	0.00017	7.11E-05	2.78E-05	1.12E-05	4.53E-06	1.81E-06
	One-time negative shock to the factor in year 2008						
A+	0.0189	0.0072	0.0028	0.0012	0.0005	0.0002	0.0001
A	0.0258	0.0105	0.0043	0.0016	0.0007	0.0003	0.0001
B+	0.0141	0.0057	0.0023	0.0009	0.0004	0.0001	0.0001
B	0.2236	0.0775	0.0239	0.0143	0.0058	0.0017	0.0007
C	0.0545	0.0214	0.0083	0.0035	0.0014	0.0005	0.0002
D	-0.00044	-0.00018	-7.10E-05	-2.78E-05	-1.12E-05	-4.53E-06	-1.81E-06

The one-time positive and negative shocks to the systematic factor have more effect on firms rated "B" at the time the shocks are imposed.

Next, we impose the same shocks on the systematic factor in year 2010 (Figure 2.22) which is a period after the financial crisis. We analyze the effect of shocks on the total default rates of all rating classes, and on the downgrade rates. Figure 2.24 shows the estimated total default rates and the stressed total default rates of all rating categories under both scenarios when we impose the shocks in year 2010.

The total default rates of all rating categories decrease by 0.54% in 2010, 0.28% in 2011 and 0.11% in 2012. Over the years 2013 and 2014, we observe less decrease in total default rates of all rating categories.

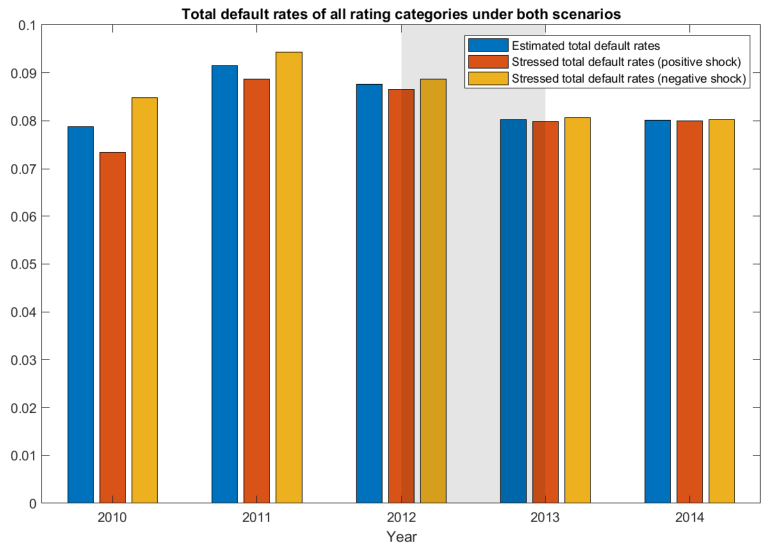


Figure 2.24: Total Default Rates of All Rating Classes Under Both Scenarios (2010). This Figure shows the estimated total default rates of all rating categories (blue) for the years 2010 to 2014, the stressed total default rates with positive shock (red) and the stressed total default rates with negative shock (yellow). The shaded areas indicate the OECD recessions in France.

Moreover, we observe an increase in the total default rates of all rating categories when we impose a negative shock to the factor. More specifically, the total default rates increase by 0.61%, 0.29% and 0.11% over the years 2010 to 2012, respectively. We also have increases of 0.04% and of 0.02% in total default rates for the years 2013 and 2014. The effect of the negative shock similarly to the positive shock is stronger in the first two years after the shock. Afterwards, the effect of shock diminishes.

The changes in downgrade rates under the positive or negative one-time shock to the systematic factor in 2010 are given in Table 2.30. The results of a one-time negative shock to the systematic factor in 2010, are similar to those of a one-time negative shock in 2008. However the size of these effects is different. The effects of one-time positive and negative shocks in the year 2008 are larger since the economy experiences recession over the period 2008 to 2009, while in year 2010, there is no recession.

Table 2.30: The Changes in Downgrade Rates

Rating class	**changes in the downgrade rates**				
	One-time positive shock to the factor in year 2010				
	2010	**2011**	**2012**	**2013**	**2014**
A+	-0.0170	-0.0076	-0.0030	-0.0011	-0.0005
A	-0.0269	-0.0103	-0.0042	-0.0017	-0.0007
B+	-0.0144	-0.0056	-0.0023	-0.0009	-0.0004
B	-0.1051	-0.0905	-0.0356	-0.0102	-0.0041
C	-0.0507	-0.0219	-0.0087	-0.0034	-0.0013
D	0.00044	0.00017	6.99E-05	2.83E-05	1.13E-05
	One-time negative shock to the factor in year 2010				
A+	0.0179	0.0078	0.0030	0.0011	0.0005
A	0.0264	0.0102	0.0041	0.0017	0.0007
B+	0.0143	0.0056	0.0023	0.0009	0.0004
B	0.1791	0.0572	0.0361	0.1050	0.0042
C	0.0530	0.0219	0.0087	0.0034	0.0013
D	-0.00044	-0.00017	-6.98E-05	-2.83E-05	-1.13E-05

We observe a decline in downgrade rates for all the rating classes except for firms rated "D" in scenario 1 (positive shock) and vice versa in scenario 2 (negative shock). Next section analyzes the effects of one-time positive and negative shocks to the systematic factor under the Ceteris Paribus assumption.

2.6.2 Shocks to the Systematic Factor- Assuming "ceteris paribus"

Under the Ceteris Paribus assumption, the factor has a flat baseline to which it returns after the shocks. We examine the effect of shocks to observe how quickly the shocks and their effects dissipate. The one-time positive and negative shocks are introduced in the year

2008, which is during a financial crisis. Alternatively, the shocks are imposed in year 2010 after the recession. Figures 2.25 and 2.26 show the paths of the systematic factor and of the stressed factor under both stress scenarios in years 2008 and 2010 respectively under the Ceteris Paribus assumption.

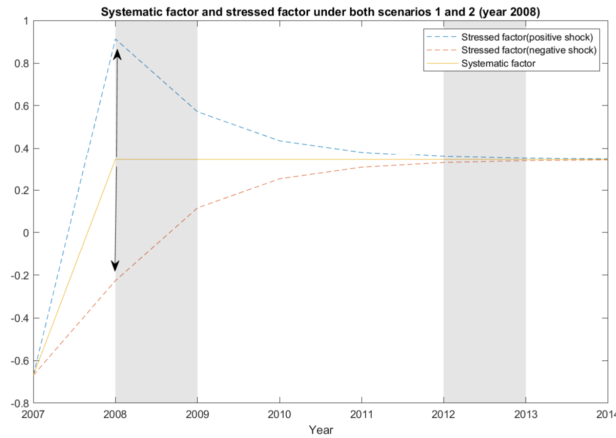


Figure 2.25: Systematic Factor and Stresses Factor Under Scenarios 1 and 2 (Year 2008). This Figure exhibits the systematic factor and the stressed factor under both scenarios, Ceteris Paribus, shocks are applied in 2008. The shaded areas indicate the OECD recessions in France.

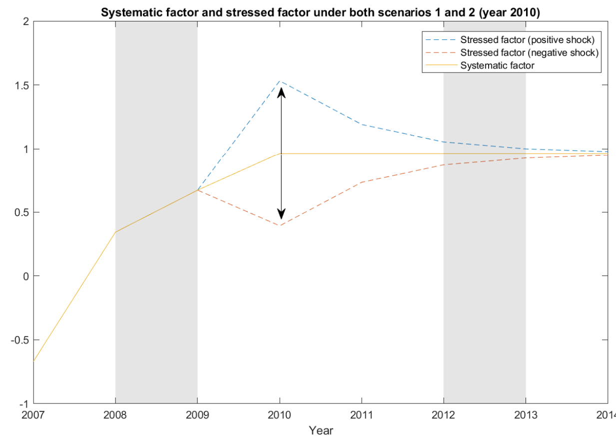


Figure 2.26: Systematic Factor and Stresses Factor Under Scenarios 1 and 2 (Year 2010). This Figure exhibits the systematic factor and the stressed factor under both scenarios, Ceteris Paribus, shocks are applied in 2010. The shaded areas indicate the OECD recessions in France.

From the graph, we observe that the effects of positive and negative shocks to the systematic factor during the financial crisis of 2008, are higher than the effects of shocks in 2010. Moreover, these effects are much stronger in the first 2 years after the shocks to the systematic factor in both years 2008 and 2010.

Next, let us compare the total default rates of all rating categories under both scenarios when we impose the shocks on the systematic factor in 2008, during the recession. The total default rates of all rating classes under both stress scenarios are shown in Figure 2.27 for year 2008 and Figure 2.28 for year 2010.

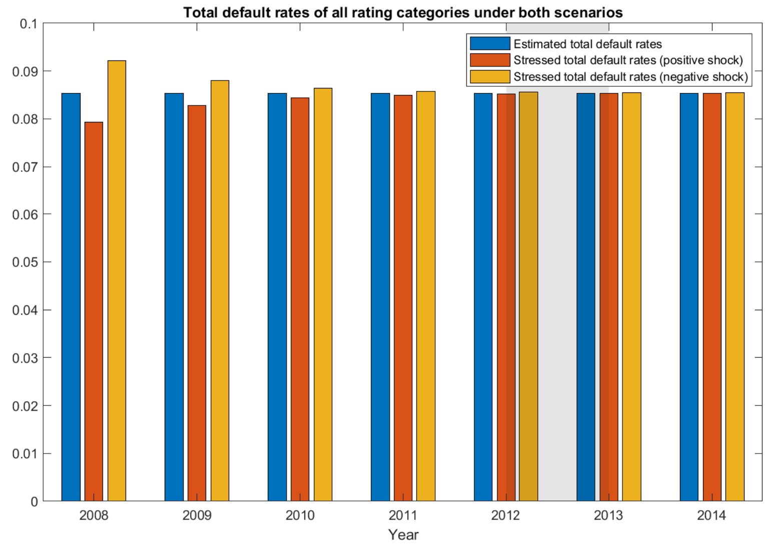


Figure 2.27: Total Default Rates of All Rating Classes Under Both Scenarios (2008). This Figure shows the estimated total default rates of all rating categories (blue) for years 2008 to 2014, the stressed total default rates after positive shock (red) and the stressed total default rates after negative shock (yellow). The shaded areas indicate the OECD recessions in France.

The total default rates decrease after a one-time positive shock to the systematic factor in year 2008. In 2008, the total default rates have decreased by 0.61% after the one-time positive shock to the factor. They have increased by 0.68% after the one-time negative shock. We observe similar dissipation patterns of both shocks in the following periods. The size of effects diminishes over time. The effect of a one-time negative shock is slightly higher than that of a one-time positive shock to the factor.

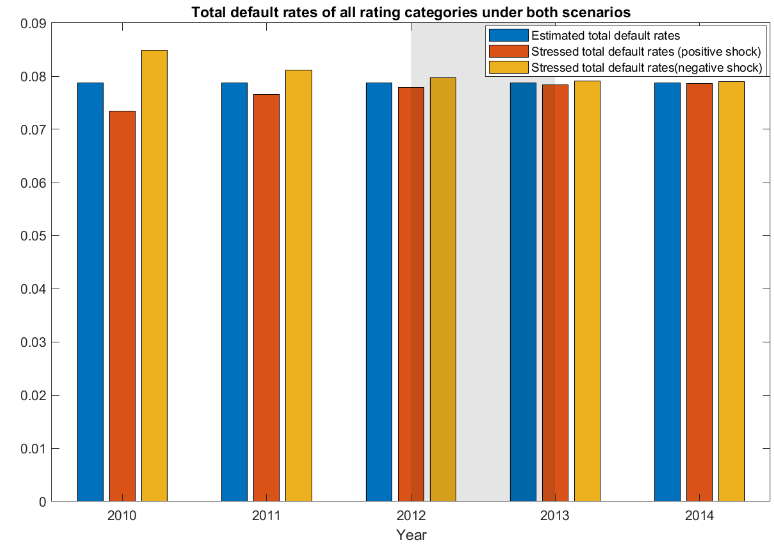


Figure 2.28: Total Default Rates of All Rating Classes Under Both Scenarios (2010). This Figure shows the estimated total default rates of all rating categories (blue) for years 2010 to 2014, the stressed total default rates after positive shock (red) and the stressed total default rates after negative shock (yellow). The shaded areas indicate the OECD recessions in France.

When the shocks are imposed on the systematic factor in years 2008, and 2010, we observe that the total defaults of all rating classes decrease after the positive shock and increase after the negative shock to the factor. The effects of shocks are stronger over the years 2010 until 2012 and diminish later on. The total default rates of all rating categories decrease by 0.54% after the positive shock and increase by 0.61% after the negative shock to the factor in 2010. This is less than the decrease (increase) observed when we imposed shocks during the recession in 2008.

Next, we explore the effects of shocks on the downgrade rates. Table 2.31 displays the changes in the downgrade rates for all rating categories when the positive and negative shocks are imposed on the systematic factor in year 2008. The one-time positive shock to the factor reduces the downgrade rates of all firms except for firms rated “D”. However, the change in rating category “D” is very small. In contrast, the one-time positive shock to the factor increases the downgrade rates of all firms except for firms rated “D”.

Table 2.31: The Changes in Downgrade Rates

Rating class	changes in the downgrade rates						
	One-time positive shock to the factor in year 2008						
	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
A+	-0.0180	-0.0073	-0.0029	-0.0012	-0.0005	-0.0002	-0.0001
A	-0.0264	-0.0105	-0.0042	-0.0017	-0.0007	-0.0003	-0.0001
B+	-0.0143	-0.0057	-0.0023	-0.0009	-0.0004	-0.0001	-0.0001
B	-0.1848	-0.0816	-0.0337	-0.0136	-0.0055	-0.0022	-0.0009
C	-0.0531	-0.0215	-0.0086	-0.0035	-0.0014	-0.0006	-0.0002
D	0.000441	0.00018	7.02E-05	2.81E-05	1.12E-05	4.49E-06	1.80E-06
	One-time negative shock to the factor in year 2008						
A+	0.0189	0.0074	0.0030	0.0012	0.0005	0.0002	0.0001
A	0.0258	0.0104	0.0042	0.0017	0.0007	0.0003	0.0001
B+	0.0141	0.0057	0.0023	0.0009	0.0004	0.0001	0.0001
B	0.2236	0.0882	0.0347	0.0138	0.0055	0.0022	0.0009
C	0.0545	0.0217	0.0086	0.0035	0.0014	0.0006	0.0002
D	-0.000436	-0.00017	-7.01E-05	-2.80E-05	-1.12E-05	-4.49E-06	-1.80E-06

We observe similar results when the shocks are imposed on the factor in the year 2010. The effects of shocks are slightly smaller than when the shocks are imposed on the factor during the financial crisis. These shocks have more effects on the firms rated “B” and “C”.

The importance of assumption Ceteris Paribus, or “other things being equal or held constant”, is to determine the causation. Once, we fix the factor, we are able to describe the effects of shocks on the migration probabilities only. Under both “Ceteris Paribus” and “Mutatis Mutandis” assumptions, the one-time positive shock to the factor in both years 2008 and 2010, causes less default and downgrades for majority of firms and a one-time negative shock increases the default and downgrades of firms.

Table 2.32: The Changes in Downgrade Rates

Rating class	changes in the downgrade rates				
	One-time positive shock to the factor in year 2010				
	2010	**2011**	**2012**	**2013**	**2014**
A+	-0.0170	-0.0069	-0.0028	-0.0011	-0.0004
A	-0.0269	-0.0107	-0.0043	-0.0017	-0.0007
B+	-0.0144	-0.0058	-0.0023	-0.0009	-0.0004
B	-0.1051	-0.0509	-0.0219	-0.0090	-0.0036
C	-0.0507	-0.0206	-0.0083	-0.0033	-0.0013
D	0.0004	0.0001	7.11E-05	2.84E-05	0.0000
	One-time negative shock to the factor in year 2010				
A+	0.0179	0.0071	0.0028	0.0011	0.0004
A	0.0264	0.0106	0.0043	0.0017	0.0007
B+	0.0143	0.0057	0.0023	0.0009	0.0004
B	0.1791	0.0633	0.0239	0.0093	0.0037
C	0.0530	0.0210	0.0083	0.0033	0.0013
D	-0.00044	-0.00018	-7.10E-05	-2.84E-05	0.0000

In our stress-testing analysis, the effects of shocks to the factor are slightly higher when we assume use the Ceteris Paribus assumption especially, for the firms rated "B" and "C". Under both assumptions and both stress scenarios, we observe higher rates of default and downgrades for firms after the negative shock to the factor, and lower rates of default and downgrades for firms after the positive shock to the factor.

The regression analysis in Section 4 and the stress-testing analysis, provide better interpretation of the unobserved factor from the changes in corporate credit migration rates. As discussed in Section 2.4.3, the unobserved factor seems to be close to the dynamics of several leading macro indicators, such as, the output gap and the total industrial production variables and the short-term interest rate from. Moreover, we found graphically similar patterns between the unobserved factor and the annual return on CAC 40 index in France. However, the factor interpretation is finalized following the stress testing analysis.

Based on the first stress scenario, a one-time positive shock to the factor reduces the default rates and the downgrade rates of firms. After the one-time positive shock, the economy experiences an upturn. Therefore, we observe less default and downgrades and higher stability and upgrade rates. Alternatively, based on the second stress scenario, a one-time negative shock to the factor causes more default and downgraded firms since the economy experiences a downturn. Because, the factor is negatively correlated with the output gap, the total industrial production and the short-term interest rate, a positive shock to any of these variables decreases the factor, and a negative shock increases the factor.

Therefore, the unobserved factor could not be the output gap or the total industrial production in France, since the one-time positive shock to these two variables decreases the factor. On the other hand, the unobserved factor could be the short-term interest rate, due to the fact that once the banks decrease their interest rates, the cost of borrowing from the banks decreases and as a result, firms have less default and downgrade rates. Moreover, the systematic factor comove with the annual return on CAC 40 which represents the overall level and direction of the market in France, and the factor has similar pattern, and a positive correlation with it. An increase in stock market indices improves the financial conditions of firms. This is also a leading indicator of the future economy and reflects the expectation of future profits. From the results of stress-testing analysis, it seems that the unobserved factor reflects the movement in the short-term interest rate and the annual return on CAC 40 index in France.

2.7 Conclusion

We use the latent factor ordered-Probit model which is a dynamic (non-linear) risk factor model, to study the dynamics of credit rating matrices by introducing an unobserved (latent) common factor that represents a fundamental driving process. This model is an important tool for credit rating and default probability estimation, as pointed out in the Basel II report on Credit Risk Factor Modelling. The latent factor ordered-Probit model has been estimated on the French dataset of the Credit Agricole S.A. Bank by two estimation methodologies. First, the latent factor ordered-Probit model has been written as an approximate linear state space model and estimated by means of the standard Kalman filter. The unobserved factor has been obtained exogenously by the principal component analysis in this approach.

Next, the latent factor ordered-Probit model has been estimated by the two-step efficient estimation approach and also alternatively by the joint optimization. The estimates of the approximate linear state space model have been used as initial values in the numerical algorithm to maximize the log-likelihoods in the two-step efficient estimation and the joint optimization approaches. In the two-step approach, the factor values are considered as nuisance parameters, and the estimator of the micro-component as a fixed effects estimator. In the second step the two-step procedure, the unobserved factor values have been replaced by the cross-sectional factor approximations and the estimator of the macro-component has been obtained by applying the maximum likelihood on the autoregressive AR(1) model. Moreover, we assessed the stochastic migration model on its ability to link the transition probabilities to an unobserved dynamic risk factor. Furthermore, two stress testing exercises were performed, with adding one time positive and negative shock to the factor. The shock effects have been used to evaluate stressed migration probabilities and default probabilities. Our main finding of this empirical analysis is that the latent factor ordered-Probit model fits well the dataset. The implementation of the two-step efficient estimation methodology is easy and less time consuming than the joint optimization. The estimates of the latent factor ordered-Probit model parameters obtained by the approximated linear state space model have been improved by the two-step estimator. The unobserved factor has been recovered by the two-step estimation method and it seems to be close to the dynamics of several leading macro indicators, such as, the output gap and the total industrial production variables from the national account group and the short-term interest rate from the financial markets group. Moreover, we found graphically a similarity in the pattern of the unobserved factor and the annual return on CAC 40 index in France. The factor interpretation is carried out by using a stress testing analysis. The migration probabilities have been effected by the one time positive and negative shocks to the factor. Due to the results of stress-testing analysis, it seems that the unobserved factor comove with the short-term interest rate and the annual return on CAC 40 index in France. Therefore, the model can be used for macro-stress testing purposes.

Chapter 3

Maximum Composite Likelihood Estimation of the Stochastic Migration Model

3.1 Introduction

This chapter introduces a new approach which is the maximum composite likelihood estimation method for a stochastic factor ordered-Probit model of credit rating migration matrices. The maximum composite likelihood extends the statistical literature and offers faster and more robust estimation than the full information maximum likelihood in a stochastic migration model. The composite likelihood is also known as the pseudo-likelihood. This method disregards some of the complex dependencies between the observations in the full joint density function. For instance, the joint densities can be replaced by a product of marginal densities or a product of conditional densities. Therefore, in the composite likelihood method, the objective function is derived by multiplying a collection of component likelihoods. The composite likelihood was first proposed by Lindsay (1988). The individual component likelihoods which correspond to marginal or conditional densities of single observation are multiplied together to define the likelihood-type objective function. The main purpose of the composite likelihood methodology is to reduce the computational complexity in models with large datasets and complex dependencies, in a setting where the full likelihood is difficult to construct.

This method provides consistent, and asymptotically normal estimators of the transition probabilities that can be used in forecasting and stress-testing. Banks implement stress-testing either for regulatory requirements or for internal capital allocation specially after the global financial crisis that resulted in the melt-down of the U.S. subprime mortgage market [See e.g. Sorge (2004); Basel Committee on Banking Supervision (2009); Blaschke (2001); Drehmann (2008); Bunn et al. (2005)]. The new method illustrated in a simulation study that confirms good performance of the maximum composite likelihood estimation.

This chapter is organized as follows. Section 3.2 describes the general definition and notation of the maximum composite likelihood. Section 3.3 explains how to construct the maximum composite likelihood for the latent factor ordered-Probit model for credit rating transitions matrices. The theoretical asymptotic properties of the estimators and assumptions are also provided in Section 3.4. Section 3.5 discussed the simulation study. The last section provides the conclusion.

3.2 Composite Likelihood Inference

3.2.1 Definitions and notation

Let $(f(y;\theta), y \in \mathcal{Y}, \theta \in \Theta)$ be a statistical parametric model with $\mathcal{Y} \subseteq \mathbb{R}^T$, $\Theta \subseteq \mathbb{R}^d$, $T \geq 1$ and $d \geq 1$. Consider a set of events $(\mathcal{A}_i : \mathcal{A}_i \subseteq \mathcal{F}, i \in \mathcal{I})$, where $\mathcal{I} \subseteq \mathbb{N}$ and $\mathcal{F}$ is some sigma algebra on $\mathcal{Y}$. The composite likelihood is defined as

$$\mathcal{L}_C(\theta, y) = \prod_{i \in \mathcal{I}} f(y \in \mathcal{A}_i; \theta)^{w_i},$$

where $f(y \in \mathcal{A}_i; \theta) = f(\{y_i \in \mathcal{Y} : y_i \in \mathcal{A}_i\}; \theta)$ with $y = (y_1, ..., y_n)$, while $\{w_i, i \in \mathcal{I}\}$ is a set of suitable weights. The associated log-likelihood is $\mathcal{L}_C(\theta; y) = \log\ \mathcal{L}_C(\theta; y)$. The weights can be ignored when they are equal. However, unequal weights improve the efficiency of the maximum composite likelihood estimators. Generally, by multiplying the component likelihoods, the composite likelihood objective function is derived. Each of these components is either a marginal or conditional density. Therefore, the derivative of the maximum composite likelihood objective function provides an unbiased estimating equation and has the properties of the likelihood based on a misspecified model. The components are multiplied regardless whether or not they are independent. The above definition can be used for combinations of both marginal and conditional densities, but the marginal and conditional versions of composite likelihoods are different [Varin et al. (2011)]. The use of either versions depends on the context. In both versions, one can choose a single observation given its neighbours or blocks of observations in the sample to define the components.

Composite Conditional Likelihoods:

Besag (1974, 1975) constructed the pseudo-likelihood from the product of the conditional densities of a single observation given its neighbour

$$\mathcal{L}_C(\theta, y) = \prod_{i=1}^{m} f(y_i|y_s;\theta),$$

where y_s is neighbour of y_i. By using the blocks of observations, Liang (1987) applied the composite likelihood to stratified case-control studies as follows:

$$\mathcal{L}_C(\theta, y) = \prod_{i=1}^{m-1} \prod_{s=i+1}^{m} f(y_i|y_i + y_s;\theta).$$

Hanfelt (2004b), Wang and Williamson (2005), and Fujii and Yanagimoto (2005) also used this form of composite likelihood. In the context of longitudinal studies, Hens et al. (2005a), and in the context of bioinformatics, Mardia et al. (2008) built composite likelihoods by combining the pairwise conditional densities

$$\mathcal{L}_{PC}(\theta, y) = \prod_{i=2}^{m} \prod_{s=1}^{m-1} f(y_i|y_s;\theta),$$

or by combining the full conditional densities

$$\mathcal{L}_C(\theta, y) = \prod_{i=1}^{m} f(y_i|y_{(-i)};\theta),$$

where $(-i)$ refers to the vector of all the observations but y_i.

Composite Marginal Likelihoods:

It is very common to choose the marginal composite likelihood when there is a N-dimensional data vector y. However, the choice of using a single observation or blocks of observations depends on the independence assumptions. The marginal composite likelihoods with the single observation component needs the independence assumption for the likelihood components. It is the simplest marginal composite likelihood

$$\mathcal{L}_{ind}(\theta, y) = \prod_{i=1}^{m} f(y_i; \theta).$$

In the literature, this marginal composite likelihood is called the independence likelihood and allows for inference on marginal parameters only [Chandler and Bate (2007)]. Moreover, if one needs to estimate the parameters related to dependence, the blocks of observations need to be modeled as in the pairwise composite likelihood [see, Cox and Reid (2004); Varin (2008)].

$$\mathcal{L}_{pair}(\theta, y) = \prod_{i=1}^{m-1} \prod_{s=i+1}^{m} f(y_i, y_s; \theta).$$

In a time-series framework, the pairwise composite likelihood for T independent and identically distributed observations $(y_1, \ldots, y_T)$ is as follows:

$$\mathcal{L}_{pair}(\theta, y) = \prod_{t=1}^{T-j} \prod_{j=1}^{J} f(y_t, y_{t+j}; \theta).$$

The maximum composite likelihood estimator θ_{CL} is obtained as the maximum of the log-composite-likelihood function, $\sum_{k=1}^{K} \ell_k(\theta; y) w_k$ where $\ell_k(\theta; y) = \log \mathcal{L}_k(\theta; y)$. The derivation of the asymptotic properties of the composite likelihood estimators is provided by Gourieroux and Monfort (2018). They describe the principles of the composite likelihood approaches in time-series data and explain the properties of the associated the composite likelihood estimators. Varin et al. (2011) also describe the composite likelihood estimation and the properties of the estimators in application to cross-sectional data [see e.g., Kent (1982); Lindsay (1988); Hens et al. (2005b)]. The asymptotic theory of the composite likelihood estimator is not different from the pseudo-likelihoods (or quasi-likelihoods) estimators.

Recalling the stochastic migration model discussed in Chapter 1, Section 1.5., the histories of observations (y_{it}) are not independent due to the presence of an unobserved common factor at each time that affects the transition probabilities, which become cross-sectionally dependent and also depend on the past observations $y_{i,t-1}$, as well as factors. Therefore, this chapter introduces the maximum composite likelihood estimation approach in application to a stochastic migration model to determine the dynamics and the forecasts of the credit ratings transition probabilities. In such a framework, the likelihood function involves multidimensional integrals of multi dimension, which that makes the maximum likelihood or GMM approaches intractable. The maximum composite likelihood method is a good alternative approach for the latent factor ordered-Probit model. Specifically, the maximum composite likelihood is an easier technique to estimate the stochastic migration model since, it only consists of pairwise dependencies and reduces the number of nonlinear dependencies to a level that is easy to handle. However, in the time-series framework if two pairs get more distant, the dependency between them become negligible. The choice of the lag length in an optimal way in the composite likelihood can be done by using the AIC/BIC type of criteria. However, it has to be mentioned that the theoretical results could be different in practice. When the lag length J is larger, the estimator is more efficient in theory since the results are based on the set-up with larger numbers of cross-section (N) and time-series (T) observations. But, in practice, the number of cross-section (N) and time-series (T) are finite which might reduce the efficiency after some point. A pairwise composite likelihood of pairs with at most J-lag distance apart is proposed by Tuzcuoglu (2017) to model the persistent discrete data by an AR-Probit model. In his framework, the cross-section N is large and the time-series dimension T is moderate. However, the maximum composite likelihood approach to the AR-Probit model can be used for different setups such as Large N, fixed T or Large N, large T and small N and large T. The derivatives of bivariate probabilities and the identification constraints will not be affected with different cross-section N and time-series T.

Concerning the properties of the composite likelihood estimator, there exists articles that provide asymptotic results for the properties of the composite likelihood estimators [see e.g. Lindsay (1988); Cox and Reid (2004) for the cross-sectional set-up; Zhao and Joe (2005); Gourieroux and Monfort (2018) for the time-series set-up]. The consistency and asymptotic normality of the composite likelihood estimator can be established under fairly standard regularity conditions. However, the maximum composite likelihood estimator is less efficient than the traditional maximum likelihood estimator in some contexts. In the next Section, we introduce the conditional composite likelihood for the ordered-Probit model with unobserved AR(1) factor in more details.

3.3 Composite Likelihood for Ordered-Probit Model with Unobserved AR(1) Factor

3.3.1 Transition Probabilities

The process of transition matrices $\{P_t, t = 1, ..., T\}$ consists of matrices P_t, which provide the probabilities of transitions from state l to state k between times $t-1$ and t given f. The elements of matrix P_t are:
$p_{kl,t} = p_{kl}(f_t;\theta) = \mathbb{P}[y_{i,t} = k|y_{i,t-1} = l, f_t] = \Phi(\frac{c_{k+1}-\beta_l f_t-\delta_l}{\sigma_l}) - \Phi(\frac{c_k-\beta_l f_t-\delta_l}{\sigma_l}),\ k, l = 1, ..., K.$
The product of two successive transition matrices $P_t^{(2)} = P_t P_{t-1}$ provides the probabilities of transitions at horizon 2 from state l to k between times $t-2$ and t given f. The elements of matrix $P_t^{(2)}$ depend on f_t, f_{t-1} and are given by:

$$p_{kl,t}^{(2)} = p_{kl}(f_t, f_{t-1};\theta) = \mathbb{P}[y_{i,t} = k|y_{i,t-2} = l, f_t, f_{t-1}] = \sum_{j=1}^{K}[p_{kj}(f_t,\theta)p_{jl}(f_{t-1},\theta)],$$

and can be computed from the components of matrices P_t and P_{t-1}. Let us denote by P and $P(2)$ the expectations of matrices P_t and $P_{t-1}^{(2)}$ with respect to common factor:

$$P = E(P_t), \quad P(2) = E(P_t^{(2)}) = E(P_t P_{t-1}).$$

The elements of matrix P are:

$$P = [p_{kl}] = [p_{kl}(\theta)] = E_{f_t}[p_{kl}(f_t, \theta)],$$

and are obtained by integrating out the unobserved factor value f_t:

$$\begin{aligned}
p_{kl}(\theta) &= \mathbb{P}[y_{i,t} = k | y_{i,t-1} = l] \\
&= \mathbb{P}[c_k < y^*_{i,t} < c_{k+1} | y_{i,t-1} = l] \\
&= \Phi\left(\frac{c_{k+1} - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2}}\right) - \Phi\left(\frac{c_k - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2}}\right).
\end{aligned}$$

The detail derivation of the above result is in Appendix E.1.

The elements of matrix $P(2)$ are:

$$P(2) = [p_{kl}^{(2)}] = [p_{kl}(2; \theta, \rho)] = E_{f_t, f_{t-1}}\left[\sum_{j=1}^{K} p_{kj}(f_t, \theta) p_{j,l}(f_{t-1}, \theta)\right],$$

$$\text{or} \quad p_{kl}(2; \theta, \rho) = \mathbb{P}[c_k < y^*_{i,t} < c_{k+1} | y_{i,t-2} = l],$$

where

$$\begin{aligned}
p_{kl}(2; \theta, \rho) = \int \sum_{j=1}^{K} \Bigg[\Bigg[&\Phi\left(\frac{c_{k+1} - \delta_j - \beta_j \rho f}{\sqrt{\sigma_j^2 + \beta_j^2(1-\rho^2)}}\right) - \Phi\left(\frac{c_k - \delta_j - \beta_j \rho f}{\sqrt{\sigma_j^2 + \beta_j^2(1-\rho^2)}}\right)\Bigg] \\
&* \left[\Phi\left(\frac{c_{j+1} - \delta_l - \beta_l f}{\sigma_l}\right) - \Phi\left(\frac{c_j - \delta_l - \beta_l f}{\sigma_l}\right)\right]\Bigg] * \phi(f) df.
\end{aligned}$$

The detail derivation of the above result is in Appendix E.2.

The transitions at horizon 2 involve one-dimensional integrals only and are easy to compute numerically.

3.3.2 Conditional Composite Likelihood Functions

This section presents the conditional composite likelihood function for the migration model with unobserved AR(1) factor model.

i) **The Conditional Composite Log-Likelihood at lag 1:**

The conditional composite log-likelihood function is defined as follows:

$$\begin{aligned} L_{cc}(\theta) &= \sum_{i=1}^{n}\sum_{t=2}^{T} \log\ l(y_{i,t}|y_{i,t-1};\theta), \\ &= \sum_{i=1}^{n}\sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{l=1}^{K} 1_{(y_{i,t}=k,y_{i,t-1}=l)}\ \log\ p_{kl}(\theta), \\ &= \sum_{k=1}^{K}\sum_{l=1}^{K} n_{kl}\ \log\ p_{kl}(\theta), \end{aligned} \tag{3.3.2.1}$$

where $n_{kl} = \sum_{i=1}^{n}\sum_{t=2}^{T} 1_{(y_{i,t}=k,y_{i,t-1}=l)} = \sum_{t=2}^{T} n_{kl,t}$ is the number of transitions from l to k in one step over the sampling period. The log-likelihood L_{cc} is calculated as if the observed ratings $(y_{i,t})$, $i = 1,...,n$, were independent among the individuals, while in reality they are linked by the common factor. Moreover, L_{cc} considers the rating processes $(y_{i,t})$, $i = 1,...,n$ as if these were Markov chains with transition matrix P, while $(y_{i,t})$, $i = 1,...,n$ are not Markov due to the factor integration that increases the memory of the process. Function $L_{cc}(\theta)$ depends on the parameter θ only and can not be used to obtain the factor dependencies and the coefficient ρ. This is why it is necessary to increase the lag.

ii) **The Conditional Composite Log-Likelihood at Lag 2**

The conditional composite log-likelihood at lag 2, $L_{cc,2}(\theta,\rho)$, is written as:

$$\begin{aligned} L_{cc,2}(\theta,\rho) &= \sum_{i=1}^{n}\sum_{t=2}^{T} \log\ l(y_{i,t}|y_{i,t-2};\theta,\rho), \\ &= \sum_{i=1}^{n}\sum_{t=2}^{T}\sum_{k=1}^{K}\sum_{l=1}^{K} [1_{(y_{i,t}=k)}1_{(y_{i,t-2}=l)}\ \log\ p_{kl}(1;\theta,\rho)], \\ &= \sum_{k=1}^{K}\sum_{l=1}^{K} n_{kl}^{(2)}\ \log\ p_{kl}(2;\theta,\rho), \end{aligned} \tag{3.3.2.2}$$

where $n_{kl}^{(2)} = \sum_{t=2}^{T} n_{kl,t}^{(2)}$ is the number of transitions from state l to k in two steps, with

$$n_{kl,t}^{(2)} = \sum_{j=1}^{K} (\hat{p}_{kj,t}\ n_{jl,t-1}).$$

The log-likelihood function $L_{cc,2}(\theta, \rho)$ is computed from the density of $(y_{i,t})$ conditional on $(y_{i,t-2})$ as if the rating histories $(y_{i,t})$ were cross-sectionally independent from one another and $(y_{i,t-2})$ contained all information about the past. Therefore, this is a quasi log-likelihood.

An important difference between L_{cc} and $L_{cc,2}$ is the set of parameters involved. One can expect to identify θ from L_{cc}, but not the parameter characterizing the serial dependence in f_t. Function $L_{cc,2}$ provides additional information that is sufficient to identify ρ.

iii) **The Conditional Composite Likelihood up to Lag 2**

The conditional composite log-likelihood up to lag 2 is defined by summing the previous composite log-likelihoods at lags 1 and 2:

$$L_c(\theta, \rho) = L_{cc}(\theta) + L_{cc,2}(\theta, \rho). \tag{3.3.2.3}$$

iv) **The Complete Log-likelihood**

As explained in Chapter 1, Section 1.5, the complete log-likelihood has a complicated expression including a large-dimensional integral, the dimension of which increases with T. It is known from the granularity theory [see Gagliardini and Gourieroux (2014b)], that an estimator of θ that is asymptotically equivalent to the ML estimator, can be derived by considering the log-likelihood conditional on the factor path and by maximizing this conditional log-likelihood with respect to both parameter θ and factor f_t. This log-likelihood conditional on (f_t) is:

$$\begin{aligned} L(\theta, f_1, ..., f_T) &= \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{l=1}^{K} 1_{(y_{i,t}=k, y_{i,t-1}=l)} \ \log \ p_{kl}(f_t; \theta), \\ &= \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{l=1}^{K} n_{kl,t} \ \log \ p_{kl}(f_t; \theta). \end{aligned} \tag{3.3.2.4}$$

It resembles the composite log-likelihood L_{cc} except that in the composite log-likelihood $p_{kl}(\theta)$ has been made independent of f_t by marginalizing. It can be maximized with respect to $\theta, f_1, ..., f_T$ and provides estimators of θ and factor values.

Intuitively, $\tilde{\theta}_T$ is the composite likelihood estimator, $\tilde{\theta}_T = Argmax_\theta L_{cc}(\theta)$. Under an identification restriction, this estimator is consistent of the true value of the parameter. However, as it does not take into account the serial dependence through the factor, a loss of information results and the estimator will not be asymptotically efficient. The estimator of θ obtained by maximizing the conditional likelihood (3.3.2.4) is capturing this serial dependence by the adjustments through the "nuisance" parameters $f_1, ..., f_T$ and will reach the asymptotic efficiency by the granularity theory. By considering the conditional composite likelihood up to lag 2 that involves also the serial dependence parameter ρ, we expect to partly reduce the lack of efficiency for θ.

In the next subsection, the identification constraints, order and rank conditions for each of these conditional composite likelihoods are discussed.

3.3.3 Identification

The parameters to be identified and their respective numbers are as follows:

Table 3.1: The Number of Parameters

Parameters	Number of Parameters
$c_k,\ k = 2, ..., K$	K-1
$\delta_k,\ k = 1, ..., K$	K
$\beta_k,\ k = 1, ..., K$	K
$\sigma_k,\ k = 1, ..., K$	K
ρ	1

The total number of parameters to identify is $4K - 2$. The negative two is due to the score $y^*_{i,t}$ being defined up to an increasing function. As we have supposed, it was a linear function of factor f_t, the score $y^*_{i,t}$ is defined up to a linear affine increasing function. The intercept and slope of that linear function are not identifiable.

3.3.4 Order Conditions

In this subsection, the order conditions for each conditional composite log-likelihood are discussed.

i) **Identification under** $L_{cc}(\theta)$**:**

The identifying functions are the elements $p_{kl}(\theta)$ of matrix P. There are $K(K-1)$ of these elements. Hence, the order condition is:

$$K(K-1) \geqslant 4K-1, \quad K^2 - 5K + 1 \geqslant 0.$$

This order condition is satisfied for $K \geqslant 5$.

ii) **Identification under** $L_{cc,2}(\theta, \rho)$**:**

The identifying functions are determined by observing that the factor f varies in the expression of $p_{k,l}(2; \theta, \rho)$. These identifying functions and their respective numbers are as follows:

$$\begin{aligned}
&(1) \quad \frac{c_k - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}} \quad \textit{number}: \quad K(K-1),\\
&(2) \quad \frac{\epsilon \beta_l \rho}{\sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}} \quad \textit{number}: \quad K,\\
&(3) \quad \frac{c_k - \delta_l}{\sigma_l} \quad \textit{number}: \quad K(K-1),\\
&(4) \quad \frac{\epsilon \beta_l}{\sigma_l}, \quad \textit{number}: \quad K,
\end{aligned}$$

where $\epsilon = \pm 1$ is an unknown sign, since the distribution of u is symmetric. The total number of identifying functions is $2K(K-1) + 2K = 2K^2$. Hence, the order condition is:

$$\begin{aligned}
2K^2 \geqslant 4K - 2 \quad &\Longleftrightarrow \quad K^2 - 2K + 1 \geqslant 0\\
&\Longleftrightarrow \quad (K-1)^2 \geqslant 0.
\end{aligned}$$

The order condition holds for any K.

iii) **Identification under** $L_c = L_{cc} + L_{cc,2}$**:**

The total number of functions available is the sum of functions available for each component of the total composite log-likelihood. Therefore, the order condition is:

$$3K^2 - K \geqslant 4K - 2.$$

The order condition is satisfied for any K.

Rank Conditions

Proposition 1: Under the L_{cc} log-likelihood function and the identifying constraints $c_1 = 0, \gamma_1 = 1$, we can identify the thresholds $c_k,\ k = 1, ..., K-1$, the intercepts $\delta_l,\ l = 1, ..., K$, and the $\gamma_l = \sqrt{\beta_l^2 + \sigma_l^2},\ l = 2, ..., K$.

Proof. See Appendix F.1.

Proposition 2: Under the log-likelihood function $L_{cc,2}$ and the identifying constraints $c_1 = 0, \gamma_1 = 1$, all the parameters are identified up to the common sign ϵ for $\beta_l,\ l = 1, ..., K$.

Proof. See Appendix F.2.

In order to fix the unknown sign ϵ, an additional constraint need to be introduced, such as:

$$\beta_1 > 0.$$

The unknown sign ϵ is a problem of global identification and not of local identification. Hence, when the asymptotic properties of the estimators are derived, this inequality constraint has to be taken into account to obtain the consistency of the estimator. It has no effect on asymptotic normality. The asymptotic properties of composite log-likelihood estimators are discussed in the next Section.

3.4 Asymptotic Properties of Composite Log-likelihood Estimators

In a panel data framework, the asymptotic analysis can be performed with respect to the cross-sectional dimension n and time T that tend to infinity as follows:

$$
\begin{aligned}
&(i)\ n \to \infty,\ T\ fixed\ \ : short\ panel\ asymptotic \\
&(ii)\ n\ fixed,\ T \to \infty\ \ : the\ time\ series\ asymptotics \\
&(iii)\ Both\ n, T \to \infty\ \ :\ double\ asymptotics
\end{aligned}
$$

The double asymptotic in case (iii) has been recently developed for application to big data. It corresponds to a long panel of multi dimensional time series.

In the ordered-Probit model with an unobserved factor, the asymptotic analysis existing in the literature concerns the granularity adjusted version of the (complete) maximum likelihood method [see, Gagliardini and Gouriéroux (2013); Gagliardini and Gourieroux (2014b)]. Let us denote the maximizers of equation (3.3.2.3) by $\hat{f}_{n,t},\ t = 1, ..., T$ and $\hat{\theta}_{n,T}$ and the autoregressive coefficient estimator obtained by regressing $\hat{f}_{n,t}$ on $\hat{f}_{n,t-1}$ by $\hat{\rho}_{n,T}$. Then, we have the following results:

(i) If $n \to \infty,\ T \to \infty$

a. $\hat{\theta}_{n,T}$ is consistent of θ, asymptotically normal and converges at speed $\frac{1}{\sqrt{nT}}$.

b. $\hat{f}_{n,t}$ is consistent of f_t, asymptotically normal and converges at speed $\frac{1}{\sqrt{n}}$ for any t.

c. $\hat{\rho}_{n,T}$ is consistent of ρ, asymptotically normal and converges at speed $\frac{1}{\sqrt{T}}$.

Depending on the setup, other asymptotic results can be considered. For example:

(ii) If $n \to \infty,\ T\ fixed$

a. $\hat{\theta}_{n,T}$ is consistent of θ, asymptotically normal and converges at speed $\frac{1}{\sqrt{n}}$.

b. $\hat{f}_{n,T}$ is consistent of f_t, asymptotically normal and converges at speed $\frac{1}{\sqrt{n}}$ for any t.

c. $\hat{\rho}_T$ is inconsistent.

(iii) If n is fixed and $T \to \infty$, neither $\hat{\theta}_{n,T}$, nor $\hat{f}_{n,T}$ nor $\hat{\rho}_{n,T}$ are consistent.

It is easy to see the problem from the regression used to estimate ρ, which has measurement errors $\hat{f}_{n,t} - \hat{f}_t, t = 1, ..., T$ in the regressors, which explains the lack of consistency.

Let us now discuss the consistency and asymptotic normality of the maximum conditional composite likelihood estimators of the identifiable parameters and introduce the following assumption:

Assumption 3 i) The parameter set of (θ, ρ) is compact, and strictly included in the set $\sigma_l > 0, \ \forall l, \ \rho < 1$.

ii) The model is well-specified, i.e. the true value (θ_0, ρ_0) is in the interior of the parameter set.

(i) If $n \to \infty, \ T \ is \ fixed:$

Contrary to the granularity-approximated (complete) maximum likelihood estimators, we can not expect the maximum composite likelihood estimators to be consistent, for T fixed. This is a consequence of the cross-sectional dependence due to the common systematic factor f_t. To clarify this point, let us assume $T = 2$ and consider the maximum conditional pairwise composite likelihood estimator based on L_{cc}. For $T = 2$, the conditional composite log-likelihood is as follows:

$$L_{cc,n}(c, \delta, \gamma) = \sum_{k=1}^{K} \sum_{l=1}^{K} n_{kl,2} \ \log \ p_{kl}(c, \delta, \gamma),$$

where $n_{kl,1}$ is $n_{kl,t}$ for $t = 2$. The normalized log-likelihood tends to:

$$\lim_{n \to \infty} \frac{1}{n} L_{cc,n}(c, \delta, \gamma),$$

$$= \lim_{n \to \infty} \sum_{l=1}^{K} \left(\frac{n_l}{n} \sum_{k=1}^{K} \left[\frac{n_{kl,2}}{n_l} \ \log \ p_{kl}(c, \delta, \gamma) \right] \right),$$

where $\frac{n_l}{n}$ is the frequency of individuals in category l. Thus we have:

$$\lim_{n \to \infty} \frac{n_l}{n} = \mathbb{P}[c_l < y^*_{i,2} < c_{l+1}] = p_l(\theta_0),$$

where $p_l(\theta)$ is the marginal probability of being in rating category l evaluated at the true value θ_0 of θ. It is well-defined if the systematic factor is strictly stationary. Therefore, we get:

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}L_{cc,n}(c,\delta,\gamma) &\simeq \sum_{l=1}^{L}\left[p_l(\theta_0)\left[\sum_{k=1}^{K}(\lim_{n\to\infty}\hat{p}_{kl,2})\ \log\ p_{kl}(c,\delta,\gamma)\right]\right],\\ &= \sum_{l=1}^{L}\left[p_l(\theta_0)\left[\sum_{k=1}^{K}p_{kl}(\theta_0,f_2)\ \log\ p_{kl}(c,\delta,\gamma)\right]\right],\end{aligned}$$

where θ_0 is the true value of θ. Therefore, the pseudo-true values of c,δ,γ that are the solutions of the following optimization:

$$\theta_0^* = (c_0^*,\delta_0^*,\gamma_0^*) = \underset{c,\delta,\gamma}{\text{argmax}}\sum_{l=1}^{L}\left[p_l(\theta_0)\left[\sum_{k=1}^{K}p_{kl,2}(\theta_0,f_2)\ \log\ p_{kl}(c,\delta,\gamma)\right]\right].$$

These pseudo-true values are functions of θ_0 and f_2. Therefore, they cannot be equivalent to the true values c_0,δ_0,γ_0 that depend on θ_0 only, and do not depend on f_2.

(ii) If n is fixed and $T\to\infty$:

This is the standard multivariate times series framework. The standard results for the consistency of estimator $\hat{\theta}_T$ in state-space models apply [see e.g. Fuh et al. (2006)].

(iii) If $n\to\infty,\ T\to\infty$:

Let us now consider the double asymptotics, and, for ease of exposition, the estimators of $\alpha=(c,\delta,\gamma)$ obtained by maximizing $L_{cc,n,T}$. We have:

$$L_{cc,n,T}(c,\delta,\gamma) = \sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}n_{kl,t}\ \log\ p_{kl}(c,\delta,\gamma),$$

after the normalization, the limit of the composite log-likelihood is:

$$\begin{aligned}
&\lim_{n,T\to\infty} \frac{1}{nT}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}\Big[n_{kl,t}\ \log\ p_{kl}(c,\delta,\gamma)\Big],\\
=&\lim_{n,T\to\infty} \frac{1}{nT}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}\Big[n_{l,t}\ \hat{p}_{kl,t}\ \log\ p_{kl}(c,\delta,\gamma)\Big],\\
=&\lim_{T\to\infty} \frac{1}{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\sum_{t=2}^{T}\Big[\lim_{n\to\infty}\ (\frac{n_{l,t}}{n}\ \hat{p}_{kl,t})\ \log\ p_{kl}(c,\delta,\gamma)\Big],\\
=&\lim_{T\to\infty} \frac{1}{T}\sum_{l=1}^{K}\sum_{k=1}^{K}\sum_{t=2}^{T}\Big[\mathbb{P}[c_k<y^*_{i,t}<c_{k+1},c_l<y^*_{i,t-1}<c_{l+1}|f_t,f_{t-1}]\ \log\ p_{kl}(c,\delta,\gamma)\Big],\\
=&\sum_{l=1}^{K}\sum_{k=1}^{K}\Big[\lim_{T\to\infty}\Big[\frac{1}{T}\sum_{t=2}^{T}\mathbb{P}[c_k<y^*_{i,t}<c_{k+1},c_l<y^*_{i,t-1}<c_{l+1}|f_t,f_{t-1}]\ \log\ p_{kl}(c,\delta,\gamma)\Big]\Big],\\
=&\sum_{l=1}^{K}\sum_{k=1}^{K}\Big[\ \mathbb{P}[c_k<y^*_{i,t}<c_{k+1},c_l<y^*_{i,t-1}<c_{l+1}]\ \log\ p_{kl}(c,\delta,\gamma)\Big],\\
=&\sum_{l=1}^{K}\sum_{k=1}^{K}\Big[\ p_l(\alpha_0)\ p_{kl}(c_0,\delta_0,\gamma_0)\ \log\ p_{kl}(c,\delta,\gamma)\Big],\\
=&\sum_{l=1}^{K}\Big[\ p_l(\alpha_0)\ \Big[\sum_{k=1}^{K}p_{kl}(c_0,\delta_0,\gamma_0)\ \log\ p_{kl}(c,\delta,\gamma)\Big]\Big],
\end{aligned}$$

by the property of the Kullback-Leibler divergence measure, we know that this limiting composite likelihood is maximized iff

$$p_{kl}(c,\delta,\gamma)=p_{kl}(c_0,\delta_0.\gamma_0)\ \ \forall k,l,$$

that is, if $(c^*_0,\delta^*_0,\gamma^*_0)=(c_0,\delta_0,\gamma_0)$, by the identification results in Section 3.3.3.

This proves the consistency of the estimator of the identifiable function of the parameters under regularity conditions.

Let us consider the first-order conditions defining the estimators $\hat{\alpha}_{n,T}=(\hat{c}_{n,T},\hat{\delta}_{n,T},\hat{\gamma}_{n,T})$ and their expansions. We have:

$$\begin{aligned}
&\frac{\partial L_{cc,n,T}(\hat{\alpha}_{n,T})}{\partial\alpha}=0,\\
\iff &\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}n_{kl,t}\ \frac{\partial\ \log p_{kl}(\hat{\alpha}_{n,T})}{\partial\alpha}=0.
\end{aligned}$$

Since $\hat{\alpha}_{n,T} \simeq \alpha_0 = (c_0, \delta_0, \gamma_0)$ for $n, T \to \infty$, we can perform a Taylor expansion of the first-order conditions in the neighborhood of α_0. We get:

$$\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T} n_{kl,t} \frac{\partial \log\ p_{kl}(\alpha_0)}{\partial \alpha} + \sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{t} n_{kl,t} \frac{\partial^2 \log\ p_{kl}(\tilde{\alpha}_{n,T})}{\partial \alpha\ \partial \alpha'} (\hat{\alpha}_{n,T} - \alpha_0) \simeq 0. \quad (3.4.1)$$

where $\tilde{\alpha}_{n,T}$ is an intermediate value between $\hat{\alpha}_{n,T}$ and α_0 by assuming the parameter set for α is convex. By applying the same argument as for the uniform a.s. convergence of the log-likelihood function, we know that:

$\frac{1}{nT}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}\left[n_{kl,t}\frac{\partial^2 \log\ p_{kl}(\alpha_0)}{\partial \alpha\ \partial \alpha'}\right]$ will converge a.s. to $\sum_{k=1}^{K}\sum_{l=1}^{K}\left[p_l(\alpha_0)\ p_{kl}(\alpha_0)\frac{\partial^2 \log\ p_{kl}(\alpha_0)}{\partial\alpha\partial\alpha'}\right]$, that $\frac{1}{nT}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}\left[n_{kl,t}\frac{\partial \log p_{kl}(\alpha_0)}{\partial \alpha}\right]$ will converge a.s. to $\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T}\left[p_l(\alpha_0)p_{kl}(\alpha_0)\frac{\partial \log p_{kl}(\alpha_0)}{\partial\alpha\partial\alpha'}\right]$, and

$$\begin{aligned}
\frac{1}{n\sqrt{T}}\sum_{k=1}^{K}\sum_{l=1}^{K}\sum_{t=2}^{T} n_{kl,t}\frac{\partial \log p_{kl}(\alpha_0)}{\partial \alpha} &= \frac{1}{\sqrt{T}}\sum_{k=1}^{K}\sum_{l=1}^{K}\left[\left(\frac{1}{n}\sum_{t=2}^{T} n_{kl,t}\right)\frac{\partial \log p_{kl}(\alpha_0)}{\partial \alpha}\right] \\
&= \frac{1}{\sqrt{T}}\sum_{k=1}^{K}\sum_{l=1}^{K}\left[\left(\sum_{t=2}^{T} p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0)\right)\frac{\partial \log p_{kl}(\alpha_0)}{\partial \alpha}\right] + o_p(1)
\end{aligned}$$

Let us assume that the matrix $J_0 = \sum_{k=1}^{K}\sum_{l=1}^{K}\left[p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0)\frac{\partial^2 \log\ p_{kl}(\alpha_0)}{\partial \alpha\ \partial \alpha'}\right]$ is positive definite. Then, by normalizing the expansion (3.4.1) by $(\frac{1}{n\sqrt{T}})$, we get:

$$\begin{aligned}
\sqrt{T}(\hat{\alpha}_{n,T} - \alpha_0) &\simeq \left[-\sum_{k=1}^{K}\sum_{l=1}^{K} p_l(\alpha_0)\ p_{kl}(\alpha_0)\ \frac{\partial^2 \log\ p_{kl}(\alpha_0)}{\partial\alpha\partial\alpha'}\right]^{-1} \\
&* \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left[\left(vec(p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0))\right)' * \left(vec\left(\frac{\partial \log p_{kl}(\alpha_0)}{\partial \alpha}\right)\right)\right] + o_p(1). \qquad (3.4.2)
\end{aligned}$$

where vec denotes the vectorialization with respect to the indexes k, l and $o_p(1)$ is a negligible term in probability. The factor f_t is strictly stationary, geometrically mixing.

Thus, the same property holds for the K^2 dimensional process $vec(p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0))$, as well as for the $dim(\alpha)$-dimensional process $[vec(p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0))]'[vec(\frac{\partial \ \log \ p_{kl}(\alpha_0)}{\partial \alpha})]$. Moreover, we have:

$$
\begin{aligned}
& E_0\Big[(vec(p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0)))'(vec(\frac{\partial \ \log \ p_{kl}(\alpha_0)}{\partial \alpha}))\Big] \\
&= E_0\Big[\sum_{k=1}^{K}\sum_{l=1}^{K}(p_l(\alpha_0) \ p_{kl}(\alpha_0) \ \frac{\partial \log \ p_{kl}(\alpha_0)}{\partial \alpha})\Big], \\
&= \frac{\partial}{\partial \alpha}\Bigg[E_0\left[\sum_{k=1}^{K}\sum_{l=1}^{K} p_l(\alpha_0) \ p_{kl}(\alpha_0) \ \log \ p_{kl}(\alpha)\right]\Bigg]_{\alpha=\alpha_0} = 0,
\end{aligned}
$$

since α_0 is the solution of the asymptotic optimization. As the factor process is geometrically mixing and ergodic, the quantity

$$
\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left[vec(p_{l,t}(\alpha_0)p_{kl,t}(\alpha_0))' \ vec(\frac{\partial \log \ p_{kl}(\alpha_0)}{\partial \alpha})\right],
$$

is asymptotically normal as well as

$$
\sqrt{T}(\hat{\alpha}_{n,T} - \alpha_0),
$$

follows,

Proposition 3: Under regularity conditions extending Fuh et al. (2006), the conditional composite ML estimator $\hat{\alpha}_{n,T}$ obtained by maximizing $L_{cc,n,T}(\alpha)$ is consistent, converges to the true α_0 at speed $\frac{1}{\sqrt{T}}$ and is asymptotically normal:

$$
\sqrt{T}(\hat{\alpha}_{n,T} - \alpha_0) \sim N\Big[0, J_0^{-1}(\sum_{h=-\infty}^{\infty} I_{0h})J_0^{-1}\Big],
$$

where

$$J_0 = -\sum_{k=1}^{K}\sum_{l=1}^{K} p_{l0}\; p_{kl}(\alpha_0)\; \frac{\partial^2 \log\; p_{kl}(\alpha_0)}{\partial\alpha\partial\alpha'},$$

$$I_{oh} = \left(vec\left(\frac{\partial\; \log\; p_{kl}(\alpha_0)}{\partial\alpha}\right)\right)' Cov_0\left[vec(p_{l,t}(\alpha_0)\; p_{kl,t}(\alpha_0)), vec(p_{l,t-h}(\alpha_0)\; p_{kl,t-h}(\alpha_0))\right]$$

$$* \left(vec\left(\frac{\partial\; \log\; p_{kl}(\alpha_0)}{\partial\alpha}\right)\right),$$

$$h = 1, 2, ...,$$

As expected:

(i) The speed of convergence is $\frac{1}{\sqrt{T}}$ instead of $\frac{1}{\sqrt{nT}}$ as in the granularity approximated (complete) log-likelihood, as a consequence of the crude cross-sectional aggregation of the data, as if the observations $y_{i,t}$ were cross-sectionally independent.

(ii) The asymptotic variance is obtained from the "sandwich" formula, as it is common in a miss-specified (pseudo) maximum likelihood approach.

(iii) The terms of type $(p_{l,t}p_{kl,t})$ and $(p_{l,t-h}p_{kl,t-h})$ depend on f_t, f_{t-1} and f_{t-h}, f_{t-h-1}, respectively. They are correlated due to the factor dynamics (except for the case when $\rho = 0$ that is of an i.i.d. factor). Therefore, the covariances have to be taken into account even if we can consider only a small number of values of lag h, if the autoregressive coefficient is not too close to the unit root.

The above asymptotic analysis is different from the literature on composite likelihood that considers either i.i.d. individuals, or finite dimensional time series [see e.g. Cox and Reid (2004); Varin et al. (2011)].

3.5 Simulation Experiments

In this section, we undertake a Monte Carlo experiment to assess the small sample performance of the proposed estimators.

3.5.1 The Designs

The designs include $K = 8$ ratings, which satisfy the order conditions in Section 3.4.1. part (i). These eight states can be related to the Standard & Poor's credit ratings scale for long-term bonds, and can be summarized into the following eight categories: AAA, AA, A, BBB, BB, B, CCC/CC, and D. The best rating is AAA, which means an "extremely strong" capacity of the borrower to repay its debt, while the worst rating is D, which means the issuer of the bond is "in default". The intermediate ratings between the two extreme cases indicate a decreasing capacity to repay which corresponds to "very strong", "strong", "adequate", "faces major future uncertainties", "faces major uncertainties", and "currently vulnerable and/or has filed a bankruptcy petition", respectively. In the following, each rating is denoted by $k = 1, \ldots, 8$, where a higher k indicates a low capacity to repay debt.

The uncertainty on migrations is driven by specific shocks $u_{i,t}$ and common systematic shocks f_t. This latent factor f_t has an AR(1) structure

$$f_t = \rho f_{t-1} + \sqrt{1-\rho^2}\eta_t,\ \eta_t \sim i.i.d.N\left(0,1\right),$$

where we set the correlation parameter to be $\rho = 0.0$, 0.2, 0.4, and 0.7, and the initial value of the factor is $f_1 = 0$. Given the rating at time $t-1$, $y_{i,t-1} = l \in \{1, \ldots, 8\}$. We suppose the underlying latent continuous quantitative score $y^*_{i,t}$ can be written as

$$y^*_{i,t} = \delta_l + \beta_l f_t + \sigma_l u_{i,t},\ u_{i,t} \sim i.i.d.N\left(0,1\right),$$

where the rating is determined by:

$$y_{i,t} = k, k = 1, \ldots, 10 \Longleftrightarrow c_k \leq y^*_{i,t} < c_{k+1}, k = 1, \ldots, 8,$$

with the threshold (c_k) described in Table 3.1 and the intercepts (δ_l) described in Table 3.2.

The thresholds and intercepts are in increasing order, and their choices were driven by the need to have higher transition probabilities on the main diagonal and decreasing probabilities when a firm transits to other states. We let $\sigma_l = \beta_l = \frac{1}{\sqrt{2}}\left(1+r\right)^{l-1}$, for $l = 1, \ldots, 7$, where $r = 0.05$, such that $\gamma_1 = \sqrt{\beta_1^2 + \sigma_1^2} = 1$.

Table 3.2: Thresholds (c_k)

k	1	2	3	4	5	6	7	8	9
c_k	$-\infty$	0	1.5	3	4.5	6	7.5	9	∞

Table 3.3: Intercepts (δ_l)

l	1	2	3	4	5	6	7
δ_l	-0.5	1	2.5	4	5.5	7	8.5

Using the simulated latent scores, we determine the number of firms transiting from l to k between $t-1$ and t as $n_{kl,t}$ out of the total number of firms kept fixed, $N = 1,000$.

The corresponding expected transition matrix is computed using Lemma 1 and presented in Table 3.3.

Table 3.4: Expected Transition Matrix (in %)

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$l=1$	69.15	28.58	2.25	0.02	0.00	0.00	0.00	0.00
$l=2$	17.05	51.26	28.86	2.79	0.04	0.00	0.00	0.00
$l=3$	1.17	17.05	49.27	29.03	3.41	0.07	0.00	0.00
$l=4$	0.03	1.51	17.84	47.33	29.09	4.08	0.12	0.00
$l=5$	0.00	0.05	1.94	18.55	45.42	29.05	4.79	0.20
$l=6$	0.00	0.00	0.09	02.42	19.16	43.57	28.91	5.85
$l=7$	0.00	0.00	0.00	0.14	2.97	19.67	41.77	35.45
$l=8$	0	0	0	0	0	0	0	1

3.5.2 The Problem of the Absorbing Barrier

The last state, i.e., default, is an absorbing state. Therefore, if we follow a given population of firms, all of the firms will default at some date, or the number of non-defaulted firms will diminish. Theoretically, the process of observed ratings is asymptotically stationary with stationary distribution the point mass on default. This difficulty can be (partly) solved in two ways.

(i) One way to partly solve this issue is to treat the very risky firms before they enter into default taking into account a current change of regulation known as the resolution step. The idea is that when a firm is entering risky rating 7, the supervisory authority can take partial control of the firm and take actions to avoid default, when the default is due to transitory difficulties. We will not incorporate the resolution step in our modelling which would require modelling of the supervisory authority behavior.

(ii) An alternative approach, which we follow, consists of considering newly created firms to balance the firms entering default. When we observe that a firm goes to default, we substitute it with the new firm with the distribution of 50% in rating class AAA, 30% in AA and 20% in A. Because at the date of firm creation, which is the first date of the debt issue their rating is high, we will replace the last row of the migration matrix at the individual level,

$$0, 0, 0, 0, 0, 0, 0, 1,$$

with last row at the population level,

$$0.5, 0.3, 0.2, 0, 0, 0, 0, 0.$$

Thus, we define a matrix P^a, which is a transition matrix without the absorbing state. Then the nondegenerate stationary distribution π^a is the solution of

$$(\pi^a)' = (\pi^a)' P^a. \tag{3.5.1}$$

For each experiment, we perform $S = 1,000$ simulations of individual trajectories, with the initial ratings $y_{i,0}$ drawn in the adjusted stationary distribution conditional on the nondefault ratings. We consider different numbers of time periods, $T = 60$, $T = 120$, and $T = 240$.

3.5.3 Simulation Results

To give some insights into the accuracy of the estimated parameters, in terms of the number of time periods T and the correlation parameters ρ, we provide in Tables $3.4 - 3.7$ the mean absolute bias and the mean squared error (MSE) of the estimates. The absolute mean bias is computed by averaging the absolute value of the estimated bias over all simulations while the MSE is obtained by computing the average square deviation of the estimated parameters from their true value. Tables 3.4 and 3.5 report for each parameter its true value and the mean absolute bias of the CL(1) estimates. In Tables 3.6 and 3.7, we present for each parameter the true value and the MSE of the CL(1) estimates.

Table 3.5: Mean Absolute Bias Varying when $\rho = 0.0$ and $\rho = 0.2$

		$\rho = 0.0$			$\rho = 0.2$		
	True Values	T=60	T=120	T=240	T=60	T=120	T=240
c_3	1.50	0.10	0.07	0.05	0.10	0.07	0.05
c_4	3.00	0.24	0.17	0.13	0.24	0.17	0.12
c_5	4.50	0.42	0.31	0.23	0.42	0.31	0.22
c_6	6.00	0.66	0.48	0.35	0.65	0.47	0.34
c_7	7.50	0.94	0.68	0.50	0.92	0.67	0.48
c_8	9.00	1.30	0.92	0.67	1.20	0.91	0.65
δ_1	−0.50	0.09	0.07	0.05	0.12	0.09	0.08
δ_2	1.00	0.10	0.08	0.06	0.11	0.08	0.06
δ_3	2.50	0.21	0.15	0.11	0.21	0.15	0.12
δ_4	4.00	0.37	0.27	0.20	0.37	0.27	0.20
δ_5	5.50	0.58	0.43	0.32	0.58	0.43	0.31
δ_6	7.00	0.85	0.62	0.46	0.84	0.63	0.45
δ_7	8.50	1.20	0.85	0.62	1.20	0.86	0.62
γ_2	1.05	0.06	0.05	0.03	0.06	0.04	0.03
γ_3	1.10	0.10	0.08	0.06	0.10	0.07	0.05
γ_4	1.16	0.15	0.11	0.08	0.14	0.11	0.08
γ_5	1.22	0.20	0.14	0.10	0.19	0.14	0.10
γ_6	1.28	0.25	0.18	0.13	0.24	0.18	0.13
γ_7	1.34	0.31	0.22	0.17	0.3	0.22	0.16

Table 3.6: Mean Absolute Bias Varying when $\rho = 0.4$ and $\rho = 0.7$

		$\rho = 0.4$			$\rho = 0.7$		
	True values	T=60	T=120	T=240	T=60	T=120	T=240
c_3	1.50	0.10	0.08	0.06	0.13	0.10	0.08
c_4	3.00	0.24	0.18	0.13	0.28	0.21	0.16
c_5	4.50	0.42	0.31	0.22	0.47	0.35	0.27
c_6	6.00	0.65	0.47	0.34	0.73	0.53	0.40
c_7	7.50	0.92	0.68	0.49	1.00	0.75	0.56
c_8	9.00	1.30	0.92	0.66	1.40	1.00	0.75
δ_1	-0.50	0.18	0.17	0.17	0.34	0.36	0.36
δ_2	1.00	0.13	0.09	0.07	0.16	0.11	0.08
δ_3	2.50	0.22	0.16	0.12	0.27	0.21	0.17
δ_4	4.00	0.38	0.28	0.21	0.46	0.36	0.31
δ_5	5.50	0.60	0.44	0.33	0.71	0.56	0.46
δ_6	7.00	0.87	0.65	0.48	1.00	0.81	0.66
δ_7	8.50	1.20	0.91	0.67	1.50	1.10	0.92
γ_2	1.05	0.06	0.04	0.03	0.05	0.04	0.03
γ_3	1.10	0.10	0.07	0.05	0.09	0.07	0.05
γ_4	1.16	0.14	0.10	0.08	0.14	0.10	0.08
γ_5	1.22	0.19	0.14	0.10	0.19	0.14	0.10
γ_6	1.28	0.24	0.18	0.13	0.24	0.18	0.13
γ_7	1.34	0.30	0.22	0.15	0.30	0.22	0.16

Table 3.7: Mean Squared Error Varying when $\rho = 0.0$ and $\rho = 0.2$

		$\rho = 0.0$			$\rho = 0.2$		
	True value	$T = 60$	$T = 120$	$T = 240$	$T = 60$	$T = 120$	$T = 240$
c_3	1.50	0.02	0.01	0.00	0.02	0.01	0.01
c_4	3.00	0.09	0.05	0.03	0.09	0.05	0.02
c_5	4.50	0.29	0.15	0.08	0.27	0.15	0.08
c_6	6.00	0.70	0.36	0.19	0.66	0.36	0.18
c_7	7.50	1.50	0.74	0.39	1.40	0.74	0.37
c_8	9.00	2.70	1.40	0.72	2.60	1.40	0.69
δ_1	-0.50	0.01	0.01	0.01	0.02	0.01	0.01
δ_2	1.00	0.02	0.01	0.01	0.02	0.01	0.01
δ_3	2.50	0.07	0.04	0.02	0.07	0.04	0.02
δ_4	4.00	0.22	0.12	0.07	0.22	0.12	0.06
δ_5	5.50	0.55	0.29	0.16	0.54	0.29	0.16
δ_6	7.00	1.20	0.62	0.33	1.20	0.63	0.33
δ_7	8.50	2.30	1.20	0.62	2.20	1.20	0.62
γ_2	1.05	0.01	0.00	0.00	0.01	0.00	0.00
γ_3	1.10	0.02	0.01	0.00	0.02	0.01	0.00
γ_4	1.16	0.04	0.02	0.01	0.03	0.02	0.01
γ_5	1.22	0.07	0.03	0.02	0.06	0.03	0.02
γ_6	1.28	0.11	0.05	0.03	0.10	0.05	0.03
γ_7	1.34	0.18	0.08	0.04	0.17	0.08	0.04

Table 3.8: Mean Squared Error Varying when $\rho = 0.4$ and $\rho = 0.7$

		$\rho = 0.4$			$\rho = 0.7$		
	True Values	$T = 60$	$T = 120$	$T = 240$	$T = 60$	$T = 120$	$T = 240$
c_3	1.50	0.02	0.01	0.00	0.03	0.01	0.01
c_4	3.00	0.09	0.05	0.03	0.12	0.06	0.04
c_5	4.50	0.27	0.15	0.078	0.34	0.19	0.11
c_6	6.00	0.66	0.36	0.19	0.81	0.44	0.25
c_7	7.50	1.40	0.75	0.38	1.70	0.90	0.50
c_8	9.00	2.60	1.40	0.69	3.10	1.70	0.91
δ_1	-0.50	0.05	0.04	0.04	0.16	0.15	0.15
δ_2	1.00	0.03	0.01	0.01	0.04	0.02	0.01
δ_3	2.50	0.08	0.04	0.02	0.12	0.07	0.04
δ_4	4.00	0.22	0.13	0.07	0.33	0.20	0.13
δ_5	5.50	0.55	0.31	0.17	0.77	0.47	0.31
δ_6	7.00	1.20	0.69	0.36	1.70	1.00	0.64
δ_7	8.50	2.40	1.40	0.71	3.40	2.00	1.30
γ_2	1.05	0.01	0.00	0.00	0.00	0.00	0.00
γ_3	1.10	0.02	0.01	0.00	0.01	0.01	0.00
γ_4	1.16	0.03	0.02	0.01	0.03	0.02	0.01
γ_5	1.22	0.06	0.03	0.02	0.06	0.03	0.02
γ_6	1.28	0.10	0.05	0.02	0.11	0.05	0.02
γ_7	1.34	0.16	0.08	0.04	0.16	0.08	0.04

Inspecting the tables, the main conclusion is that for each value of the correlation parameter ρ, the CL(1) estimates have a small estimated bias and MSE. Both decrease as the number of time periods increases. For instance, when $\rho = 0.0$, the largest absolute bias is 1.3 when estimating the threshold parameter c_8 and $T = 60$, which remains small relative to its true value $c_8 = 9$. This bias decreases to 0.67 when $T = 240$. These results are consistent with the asymptotic results on the $\sqrt{T}-$consistency of the CL(1) estimates in Proposition 3. In addition, we notice that in general, there is a very small effect of ρ on the consistent estimation of the different parameters. However, the impact varies with the amount of correlation in the score induced by the latent factor changes, when conducting inferences. This aspect is discussed below.

We illustrate the finite sample performance of the inferences based on t statistics by plotting the empirical probability distribution function (PDF) of each computed t statistic when testing that each of the parameters equals their true value (Figures $3.1-3.12$). In Figures $3.1-3.4$, we plot the empirical probability distribution by depicting the histogram of the computed test statistics over S simulations for the threshold parameters $(c_3, \ldots, c_8)$ when the time dimensions change for the different values of ρ. Figures $3.5-3.8$ show the empirical probability distribution for the intercepts $(\delta_1, \ldots, \delta_7)$, while Figures $3.9-3.12$ present the distribution for the unconditional variances of the latent scores $(\gamma_2, \ldots, \gamma_7)$. To compute the standard errors, we use the diagonal elements of the estimator $\hat{\Sigma}_\theta = \widehat{Var}\left(\sqrt{T}\hat{\theta}_{nT}\right)$ of the asymptotic covariance matrix $\Sigma_\theta = J_0^{-1}\left(\sum_{h=-\infty}^{\infty} I_{0h}\right) J_0^{-1}$ in Proposition 3, which is the heteroskedasticity and autocorrelation consistent (HAC) estimator. The HAC estimator is obtained using a quadratic spectral kernel and a bandwidth, which we set to be $4\,(T/100)^{2/9}$ following Newey and West (1994). In each figure, the x-axis represents the computed t statistic over the simulations, while the y-axis contains their frequencies. For $\rho = 0.0$ and $\rho = 0.2$, a common feature in the tables is that when T varies the distribution of t statistics is centered on zero and is symmetric. In most cases, the simulated t statistics belong to the intervals $[-1.96, 1.96]$, as would be expected for asymptotically normally distributed t statistics, when the level of the test is 5%. However, when the correlation parameter becomes higher ($\rho = 0.4$ and $\rho = 0.7$), for some of the estimated parameters (for e.g., δ_l), the realized t statistics are shifted from the zero mean due to the smaller variance, suggesting an under-estimation of the true variance. The shift is negative for the negative values of the parameters and is positive for the positive values of the parameters. The results for large values of ρ may be explained by the fact that the CL(1) estimation integrates out the latent factor by imposing they are $N(0,1)$ (but dependent), and the likelihood function does not incorporate the parameter ρ, which captures the dependence in the factor dynamic. To solve this issue, one may consider the CL(2), estimation if we have strong reason to believe there is a strong dependence in the common factor. For instance, the estimated correlation parameter in Feng et al. (2008) is 0.02. Alternatively, a numerical procedure like the bootstrap method may be used for a more accurate approximation of the asymptotic distributions. Nevertheless, these simulations are computationally intensive. Therefore, the numerical procedures in this context will have to deal with the computation issue, that is common for maximum likelihood procedures which involve a large number of parameters.

Figure 3.1: Empirical PDF of t Statistic for $\hat{c}_{k+1}$ when $\rho = 0.0$

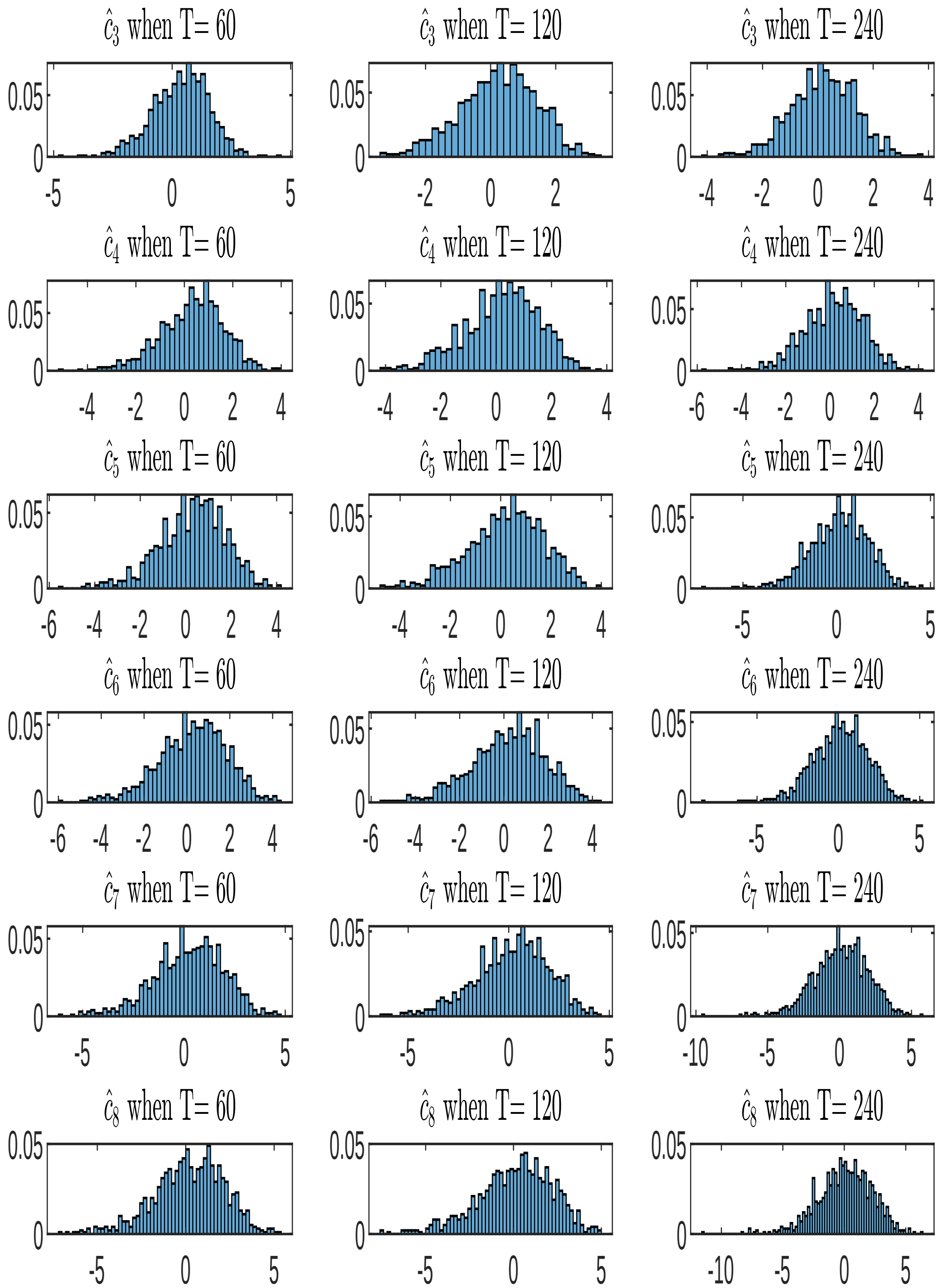


Figure 3.2: Empirical PDF of t Statistic for $\hat{c}_{k+1}$ when $\rho = 0.2$

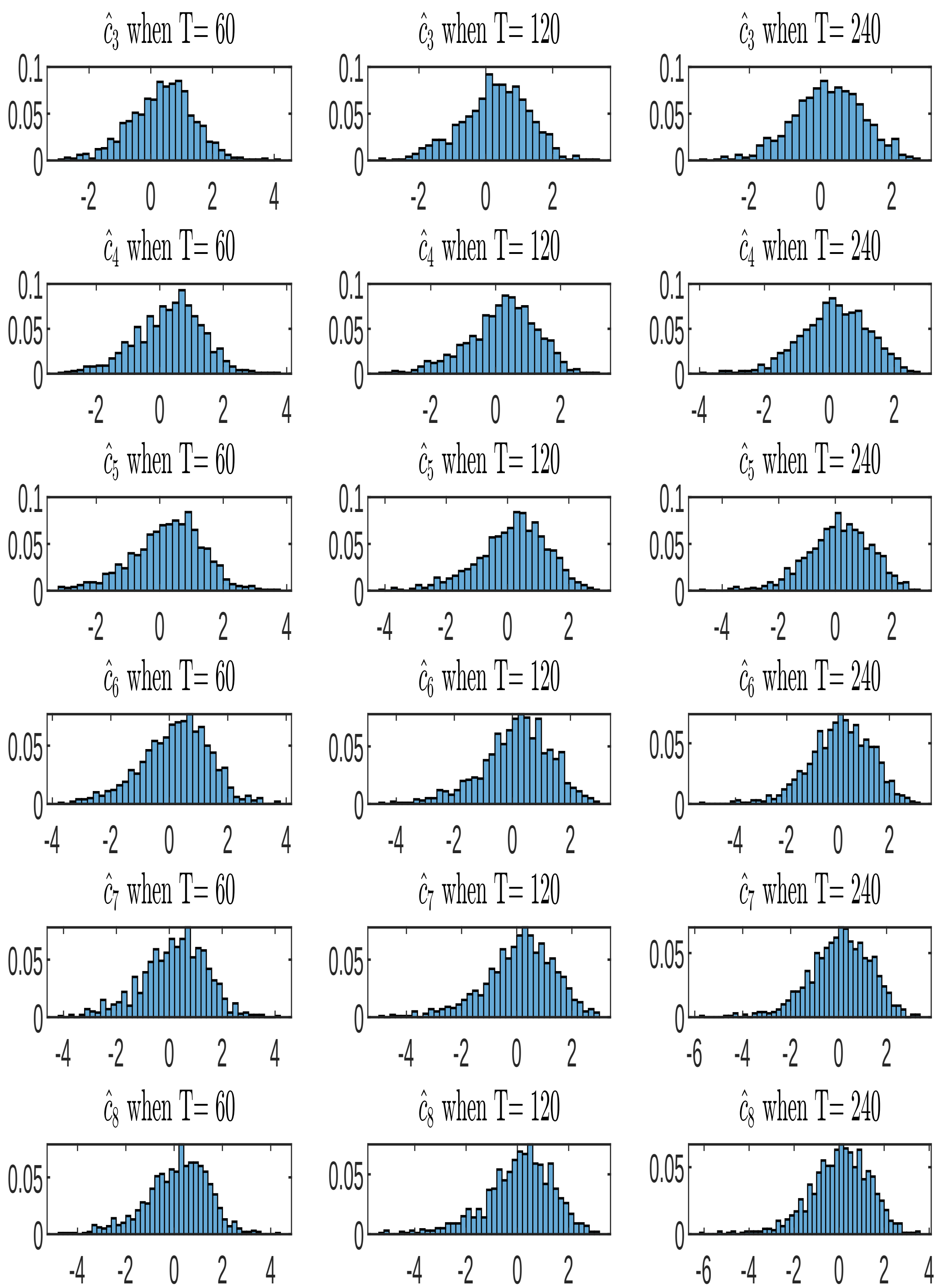


Figure 3.3: Empirical PDF of t Statistic for $\hat{c}_{k+1}$ when $\rho = 0.4$

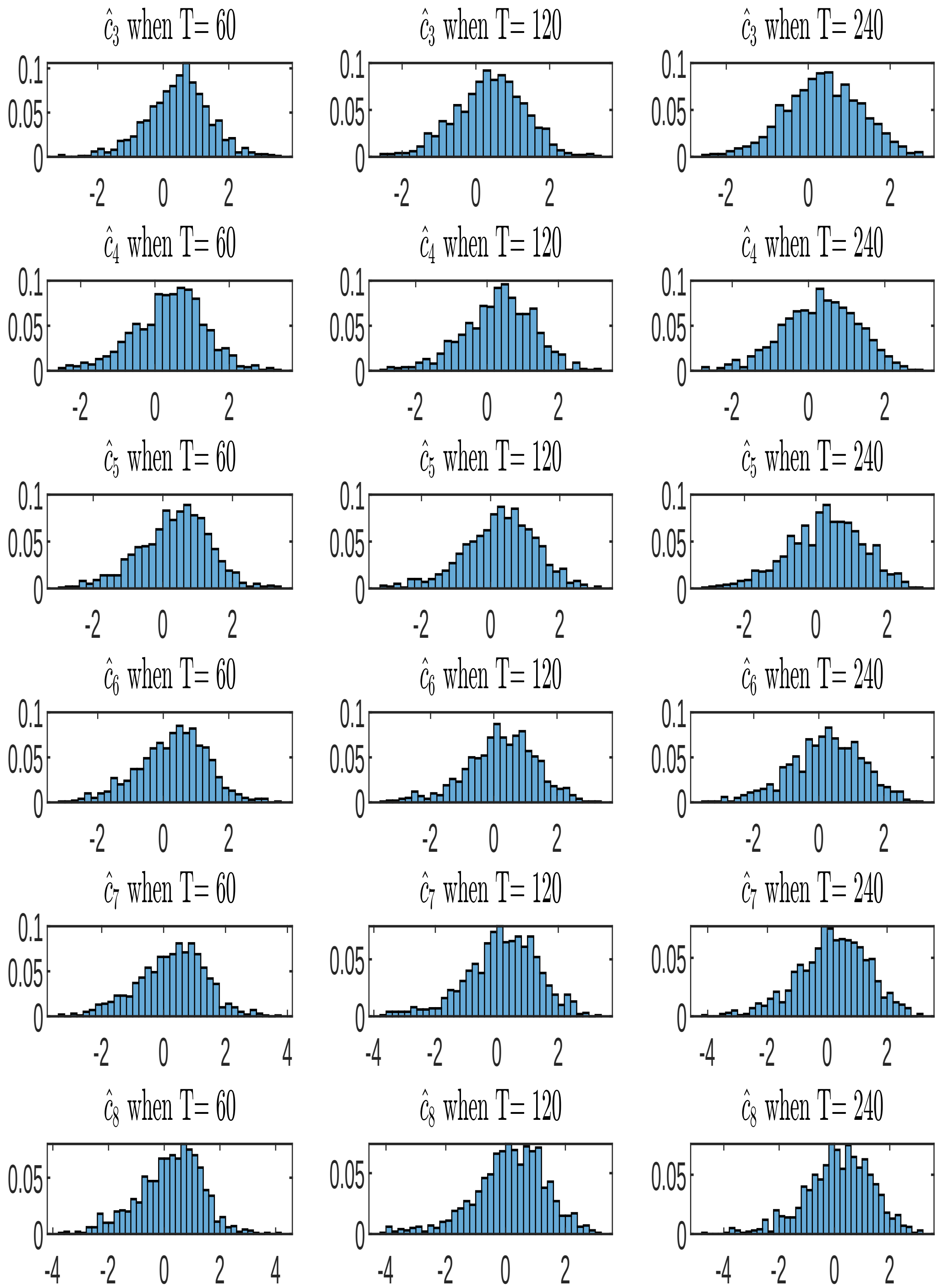


Figure 3.4: Empirical PDF of t Statistic for $\hat{c}_{k+1}$ when $\rho = 0.7$

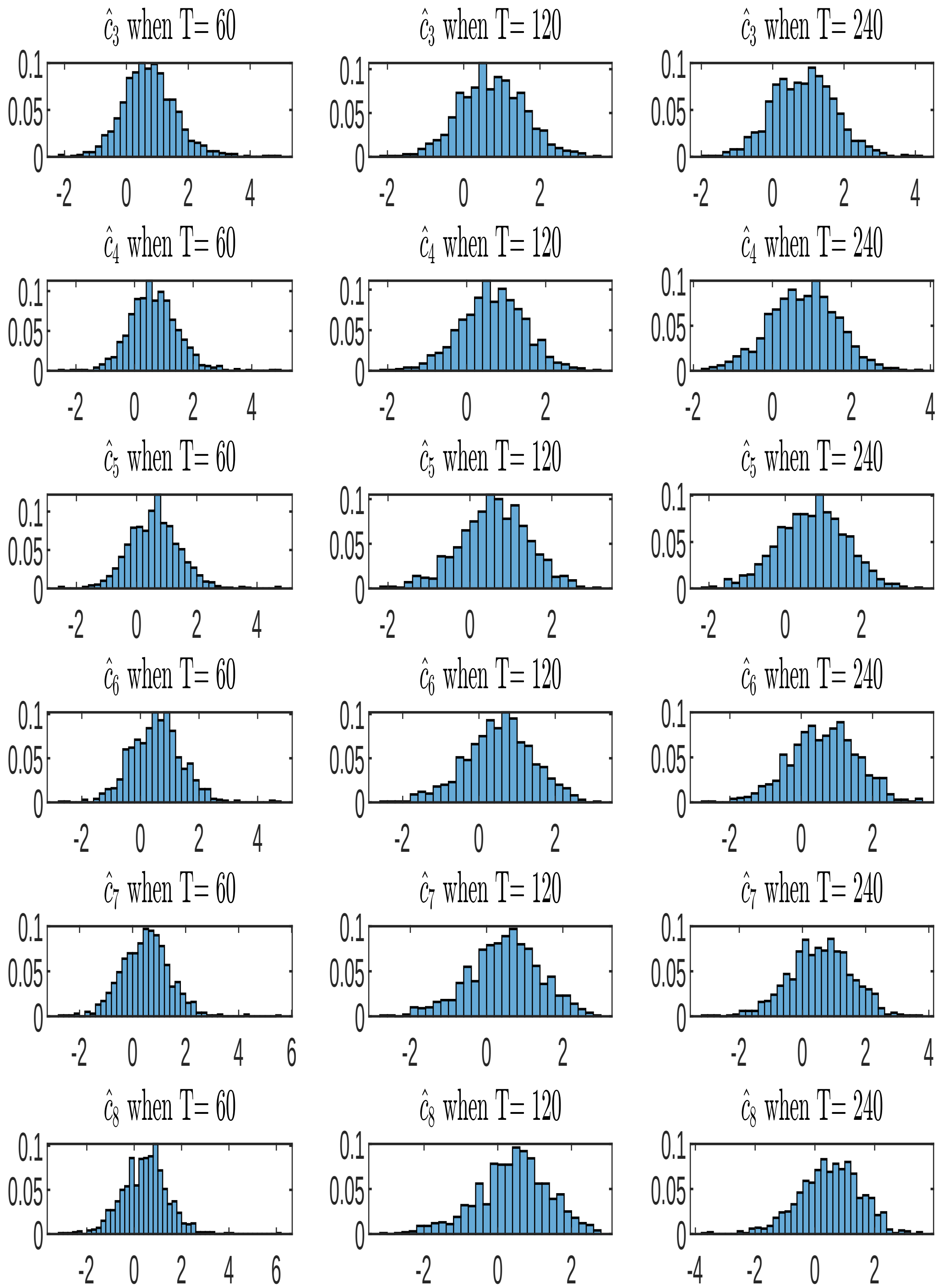


Figure 3.5: Empirical PDF of t Statistic for $\hat{\delta}_l$ when $\rho = 0.0$

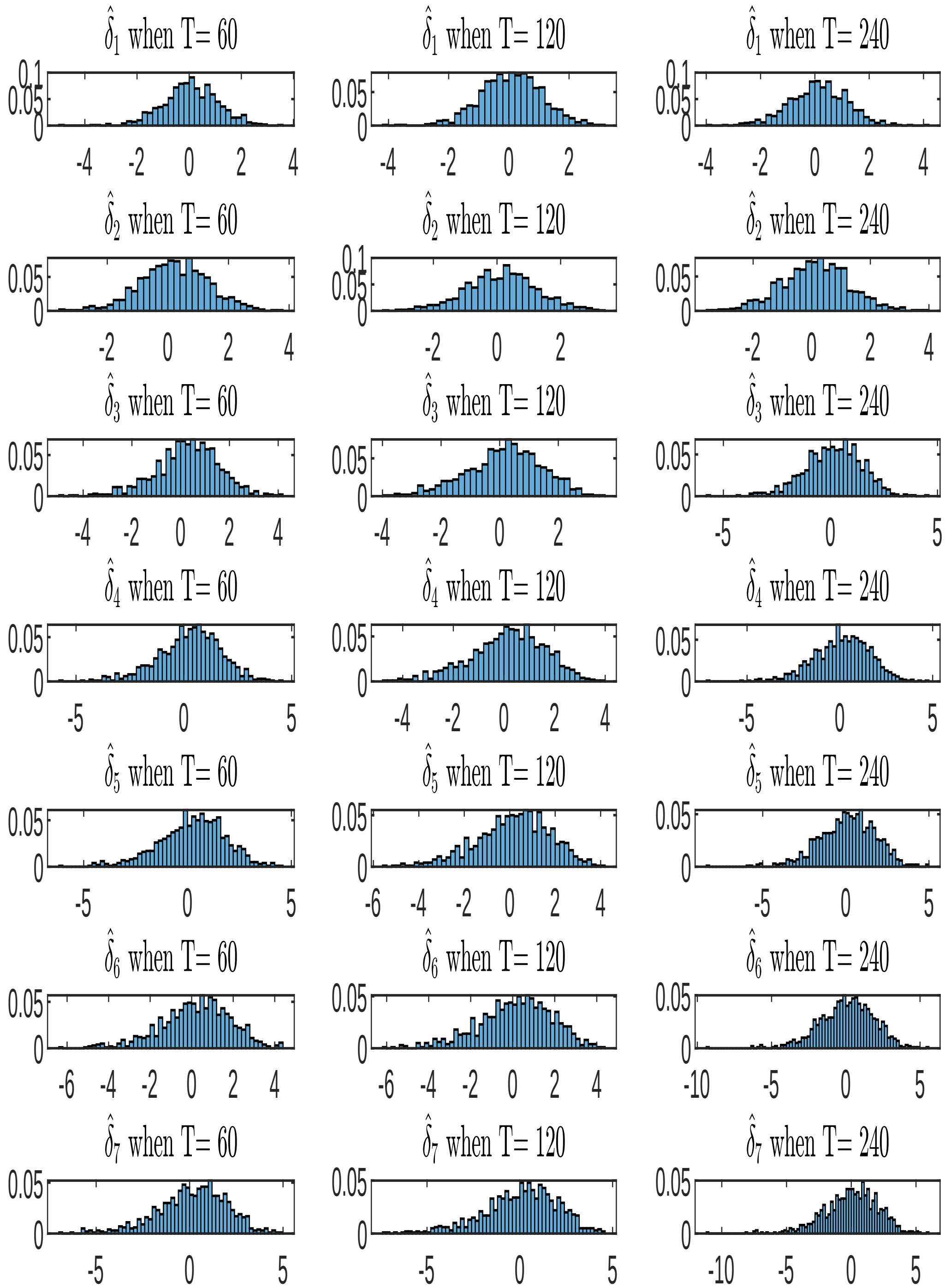


Figure 3.6: Empirical PDF of t Statistic for $\hat{\delta}_l$ when $\rho = 0.2$

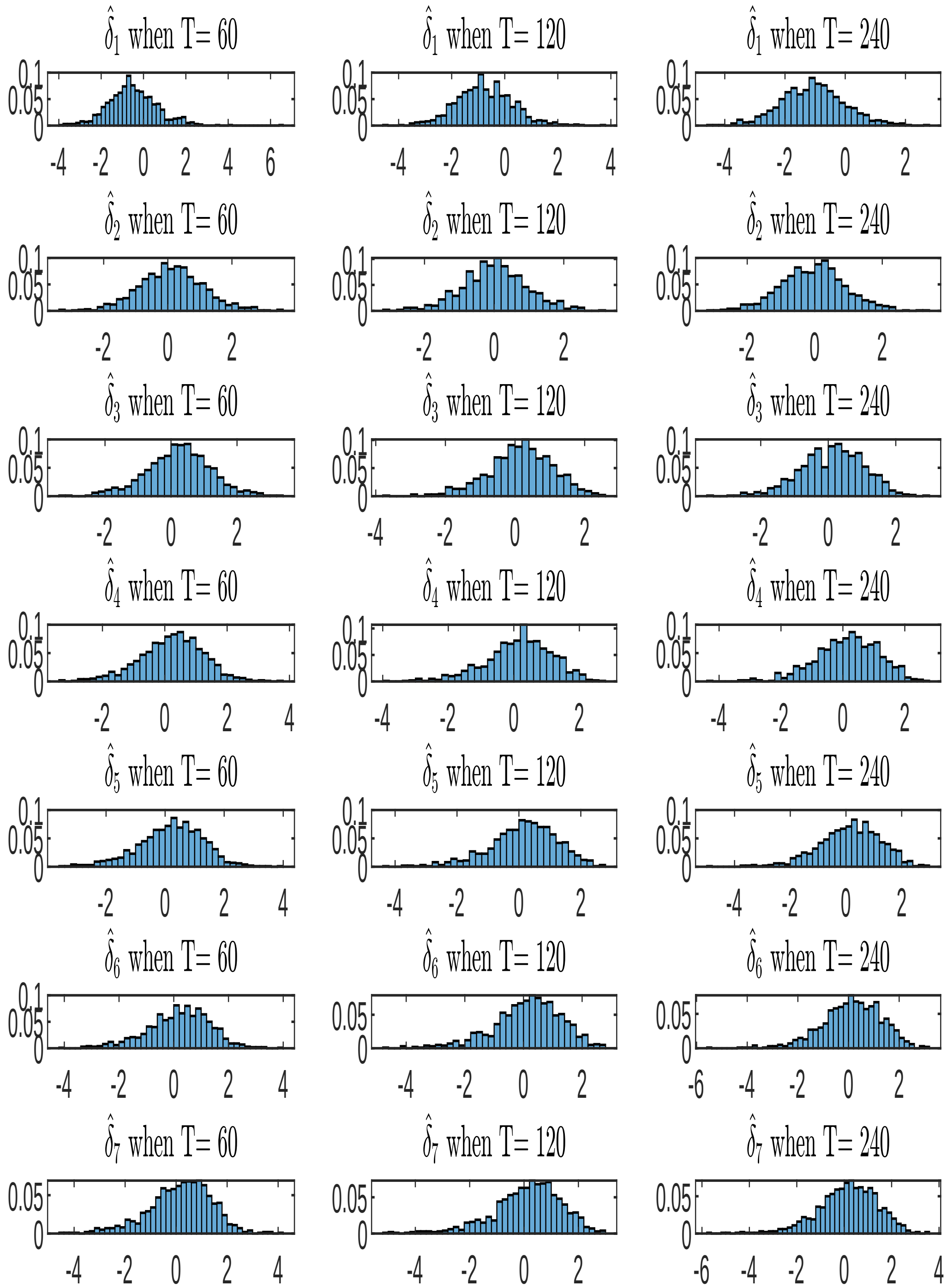


Figure 3.7: Empirical PDF of t Statistic for $\hat{\delta}_l$ when $\rho = 0.4$

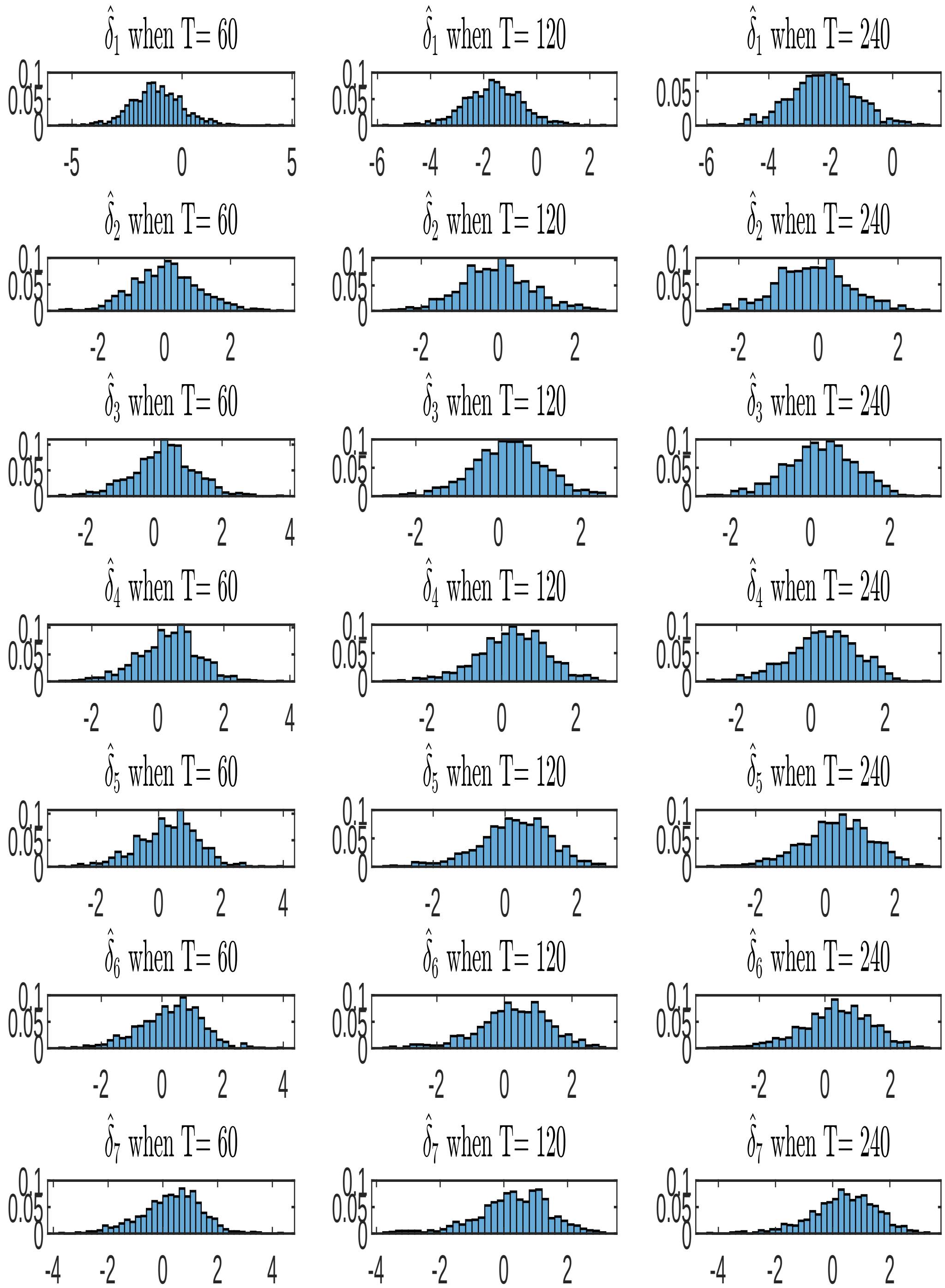


Figure 3.8: Empirical PDF of t Statistic for $\hat{\delta}_l$ when $\rho = 0.7$

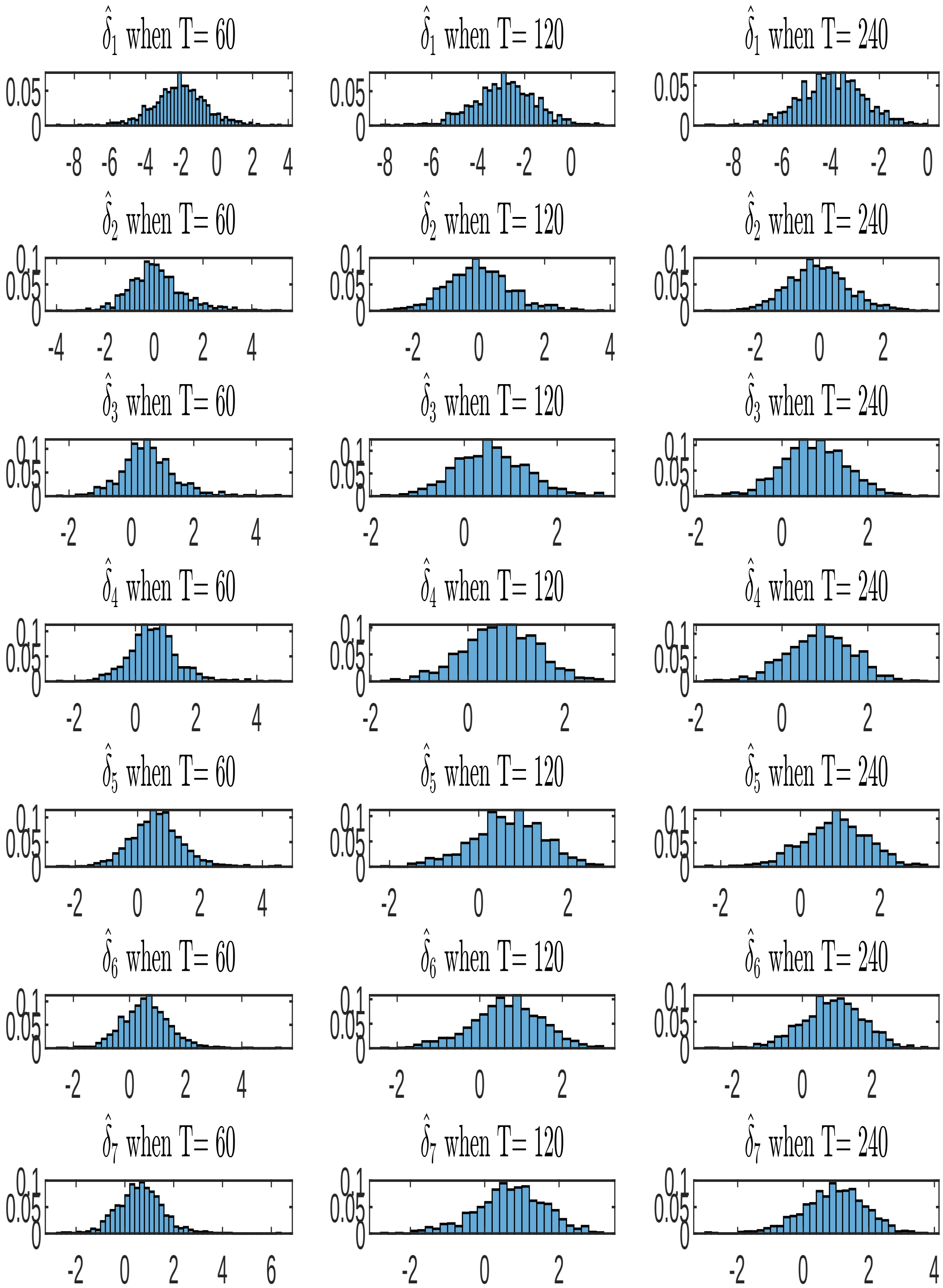


Figure 3.9: Empirical PDF of t Statistic for $\hat{\gamma}_l$ when $\rho = 0.0$

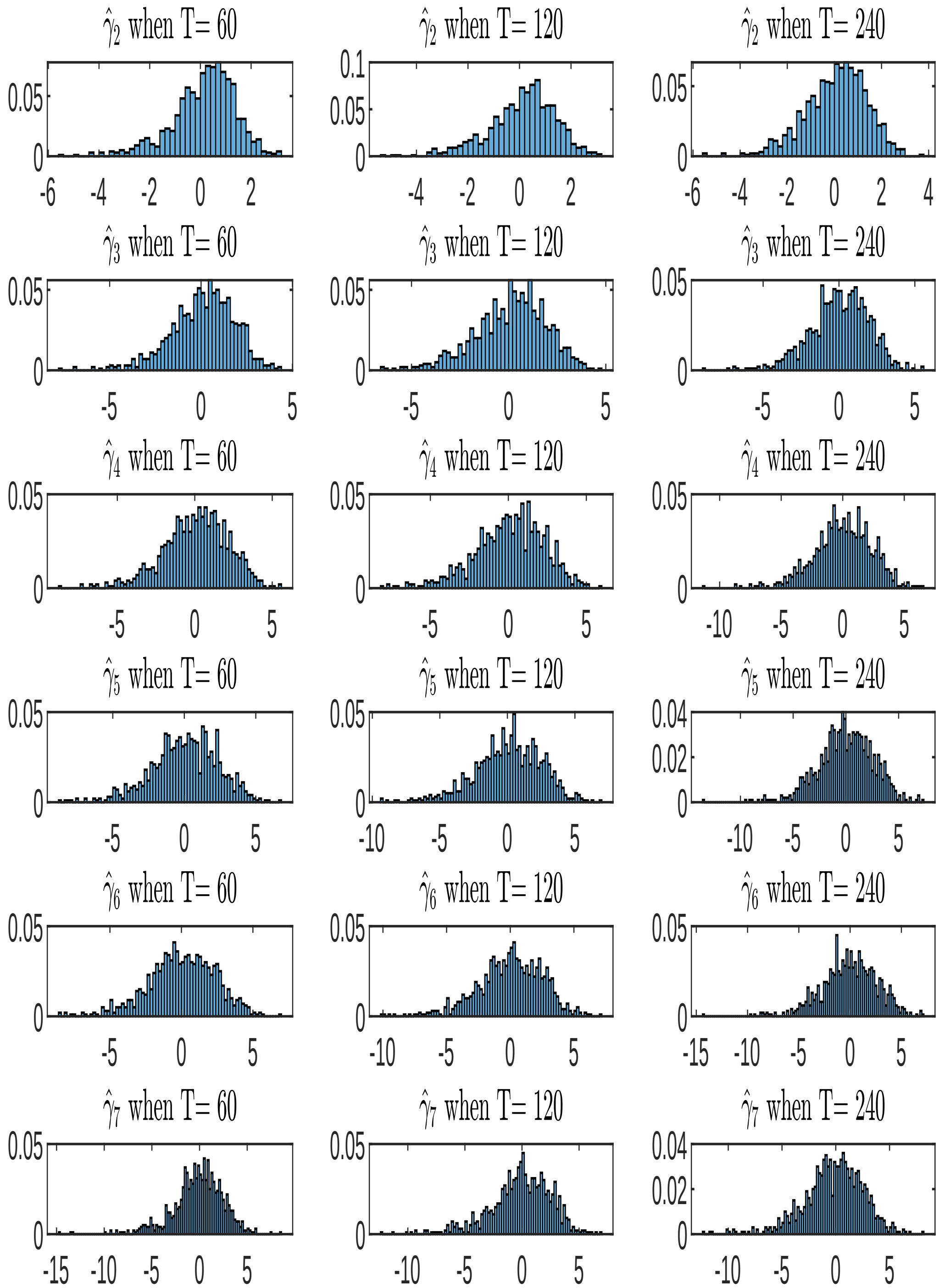


Figure 3.10: Empirical PDF of t Statistic for $\hat{\gamma}_l$ when $\rho = 0.2$

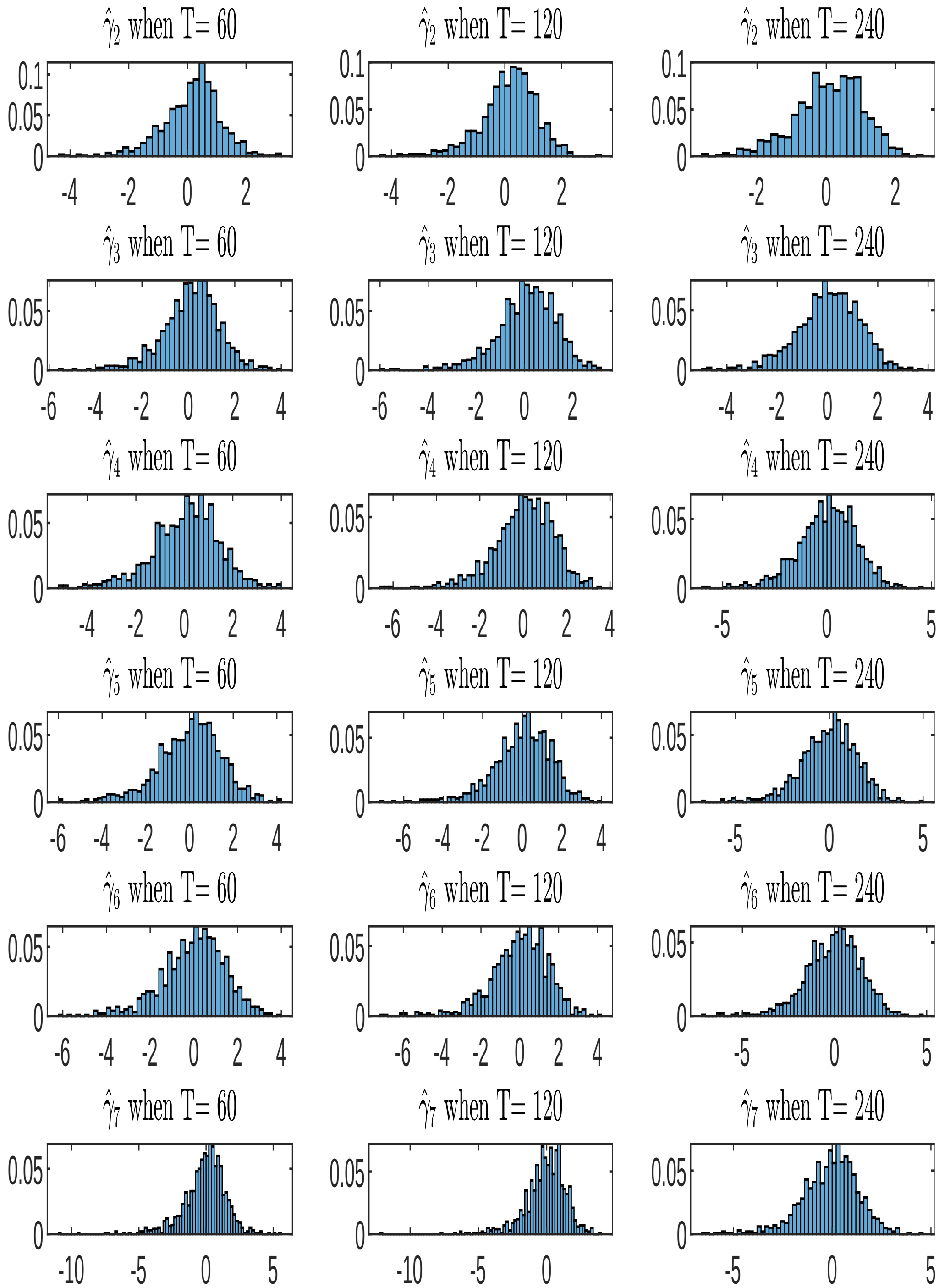


Figure 3.11: Empirical PDF of t Statistic for $\hat{\gamma}_l$ when $\rho = 0.4$

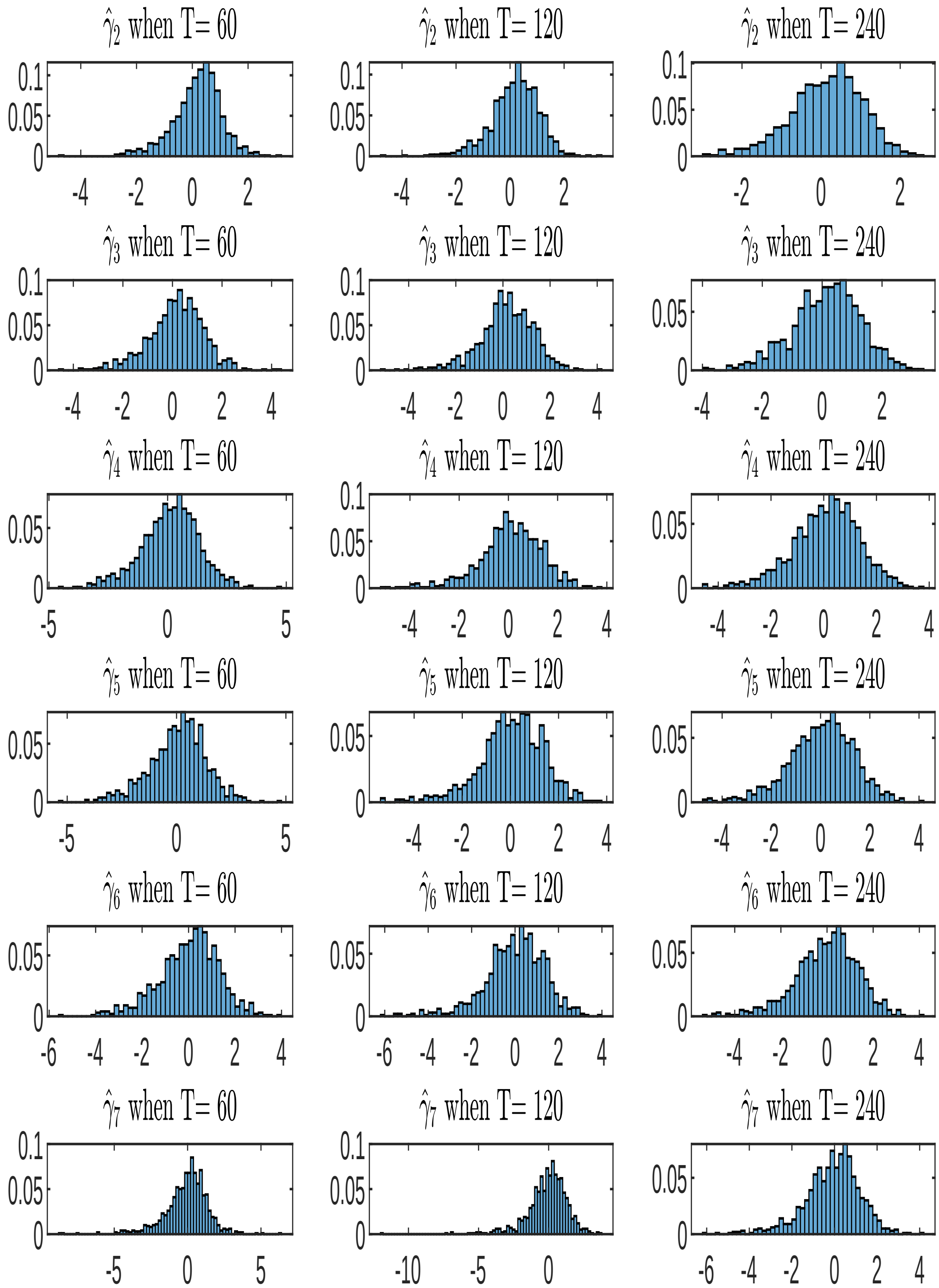
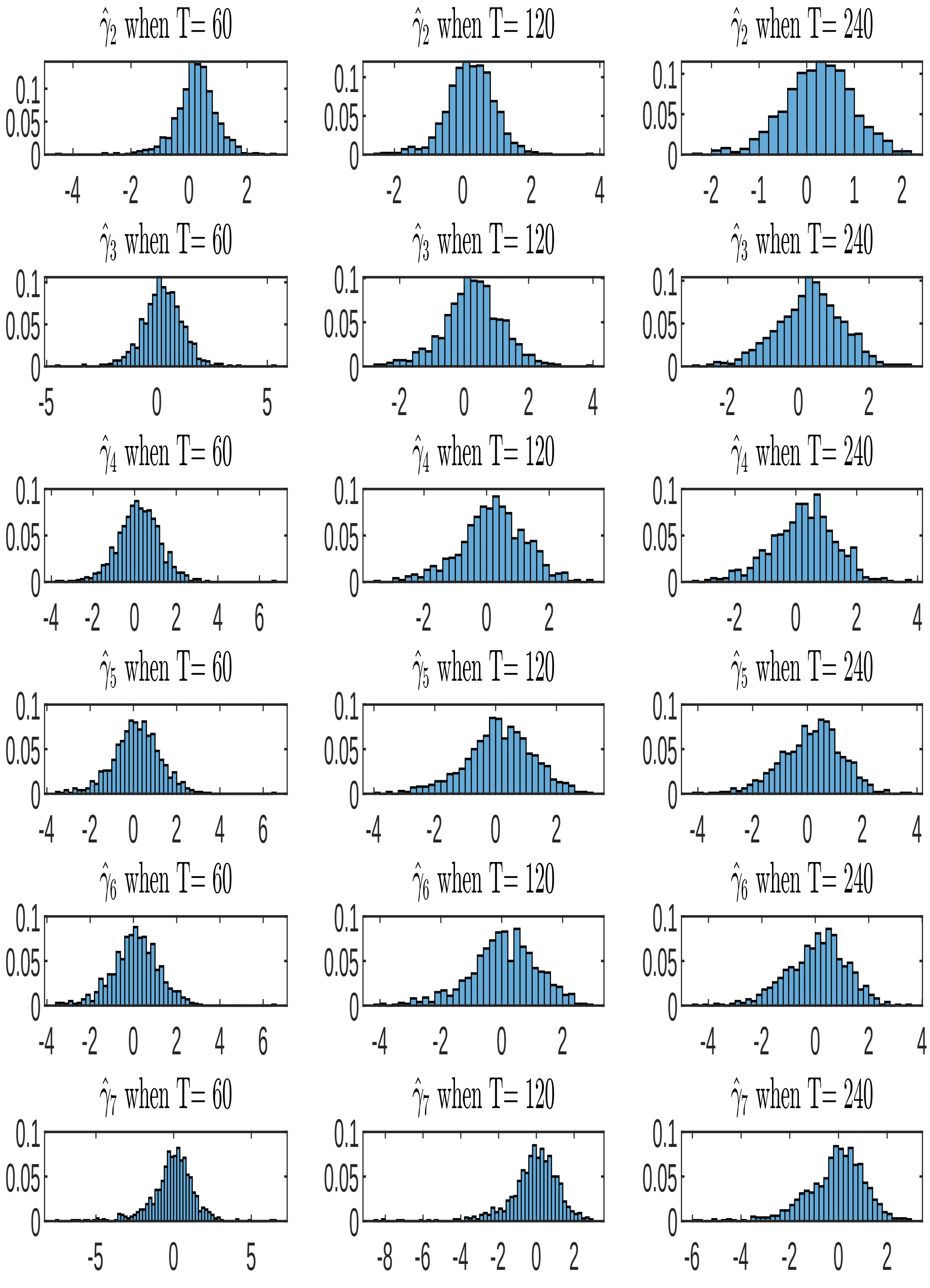


Figure 3.12: Empirical PDF of t Statistic for $\hat{\gamma}_l$ when $\rho = 0.7$



Tables 3.8 and 3.9 illustrate the ability of the CL(1) estimation to reproduce the pattern of useful parameters for downgrade risks in risk management. In particular, the first panel of Tables 3.8 and 3.9 presents the estimated downgrade probability at horizons 1 and 2 of a firm currently rated A ($l = 3$) for different numbers of time periods and different values of the correlation parameter. The expected downgrade probability for any initial rating $l = 3$ is the sum of the expected transition probabilities to $k = 4, \ldots, 8$. At horizon 1, when $T = 60$ and $\rho = 0$, the downgrade probability is 32.35%, while its true value is 32.51%. At horizon 2, when $T = 60$ and $\rho = 0$, the estimated probability increases to 43.16%, which remains close to the true value, 43.32. As noted, the downgrade probability increases as the horizon increases, and this feature is reproduced after the estimation. When the time dimension changes, this pattern remains the same. Another feature of the results is that the estimated probability at horizons 1 and 2 is in most cases more accurate when the correlation in the latent factor is small.

We now discuss the estimation of the probability of default illustrating a case where a firm is initially rated A. The second panel of Tables 3.8 and 3.9 shows the estimated expected probability of default after plugging in the estimated parameters. At horizon 1, all the probabilities remain the same for all values of T, and equal zero. The probability that a firm with a strong capacity to repay its debt (a firm rated A) defaults is negligible. When the horizon increases, this probability increases as expected. More accurate probabilities of default are obtained for small values of ρ.

Table 3.9: Downgrade Probability and Probability of Default (in %) at Different Horizon for a Firm Initially Rated A when $\rho = 0.0$ and $\rho = 0.2$

		$\rho = 0.0$			$\rho = 0.2$		
	True Value	$T = 60$	$T = 120$	$T = 240$	$T = 60$	$T = 120$	$T = 240$
$DP\,(1\|A)$	32.51	32.35	32.49	32.56	32.28	32.44	32.54
$DP\,(2\|A)$	43.32	43.16	43.30	43.42	43.09	43.28	43.45
$PD\,(1\|A)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$PD\,(12\|A)$	23.49	23.47	23.56	23.65	24.91	25.18	25.37
$PD\,(24\|A)$	55.15	53.51	54.41	54.98	54.77	56.22	57.11
$PD\,(36\|A)$	74.09	70.85	72.49	73.45	71.27	73.63	74.99
$PD\,(48\|A)$	85.07	81.01	83.02	84.14	80.83	83.60	85.13
$PD\,(60\|A)$	91.39	87.20	89.29	90.40	86.68	89.51	90.99

Table 3.10: Downgrade Probability and Probability of Default (in %) at Different Horizon for a Firm Initially Rated A when $\rho = 0.4$ and $\rho = 0.7$

		$\rho = 0.4$			$\rho = 0.7$		
	True Value	$T = 60$	$T = 120$	$T = 240$	$T = 60$	$T = 120$	$T = 240$
$DP\,(1\|A)$	32.51	32.25	32.51	32.63	32.66	33.09	33.40
$DP\,(2\|A)$	43.32	43.16	43.49	43.70	44.14	44.81	45.28
$PD\,(1\|A)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$PD\,(12\|A)$	23.49	26.97	27.49	27.89	32.10	33.75	34.91
$PD\,(24\|A)$	55.15	56.53	58.71	60.07	60.36	64.76	67.41
$PD\,(36\|A)$	74.09	71.93	75.16	77.06	73.12	78.61	81.71
$PD\,(48\|A)$	85.07	80.71	84.37	86.41	80.20	86.00	89.08
$PD\,(60\|A)$	91.39	86.12	89.81	91.74	84.64	90.37	93.21

3.6 Conclusion

This chapter introduced the maximum composite likelihood estimation method for the stochastic factor ordered-Probit model for credit rating migration matrices. This method provides consistent, and asymptotically normal estimators of the transition probabilities that can be used in forecasting and stress-testing. The main purpose of the composite likelihood methodology is to reduce the computational complexity in models with large datasets and complex dependencies, in a setting where the full likelihood is difficult to construct. We introduced three maximum conditional likelihood functions for the migration model with unobserved AR(1) factor model: the conditional composite likelihood at lag 1, $L_{cc}(\theta)$, the conditional composite likelihood at lag 2, $L_{cc,2}(\theta, \rho)$, and the conditional composite likelihood up to lag 2, which is the combination of the first two functions. Due to independence assumption on factor, the maximum composite likelihood function at lag 1, $L_{cc}(\theta)$, only depends on the parameter of the model and can not be used to obtain the factor dependencies. However, the function $L_{cc,2}(\theta, \rho)$ provides more information that is sufficient to identify the factor dependencies. We discussed the identification constraints, order and rank conditions for each of these conditional composite likelihoods. We provided the introductory results on the consistency of the estimators and their asymptotic normality. We showed that the asymptotic properties of the maximum composite likelihood estimators with respect to the cross-sectional dimension n and T tend to infinity can be achieved. Under an identification restriction the estimator of the maximum composite likelihood at lag 1 is consistent of the true value of the parameter, however, as it does not take into account the serial dependence through the factor, a loss of information results and the estimator will not be asymptotically efficient. We expect to partly increase the efficiency of the parameters by considering the conditional composite likelihood up to lag 2 that involves also the serial dependence parameter ρ. We undertook a Monte Carlo experiment to assess the small sample performance of the proposed estimators of $L_{cc}(\theta)$. We provided the mean squared error (MSE) ans the mean absolute bias of the estimates to give some insight into the accuracy of the estimated parameters. Our results showed that for each value of the correlation parameter ρ, the maximum composite likelihood at lag 1 estimates, have a small estimated bias and MSE, and it decreases as the number of time periods increases. These results are consistent with the asymptotic results on the $\sqrt{T}$consistency of the maximum composite likelihood at lag 1 estimates. In addition, we noticed that in general, there is a very small effect of ρ on the consistent estimation of the different parameters. However, the impact varies with the amount of correlation in the score induced by the latent factor changes, when conducting inferences. To address this aspect, we illustrated the finite sample performance of the inferences based on t statistics by plotting the empirical probability distribution function (PDF) of each computed t statistic when testing that each of the parameters equals their true value. Our findings reveal that when T varies the distribution of t statistics is centered on zero and is symmetric, for correlation parameter $\rho = 0.0$ and $\rho = 0.2$. However, when the correlation parameter becomes higher $\rho = 0.4$ and $\rho = 0.7$, for some of the estimated parameters (for e.g.,the intercept), the realized t statistics are shifted from the zero mean due to the smaller variance, suggesting an under-estimation of the true variance. The results for large values of ρ may be explained by the fact that the maximum composite likelihood at lag 1 estimation integrates out the latent factor by imposing they are $N(0, 1)$ (but dependent), and the likelihood function does not incorporate the parameter ρ, which captures the dependence in the factor dynamic. To solve this issue, one may consider the maximum composite likelihood up to lag 2, estimation if we have strong reason to believe there is a strong dependence in the common factor. Moreover, we illustrated the ability of the maximum composite likelihood at lag 1 estimation to reproduce the pattern of useful parameters for downgrade risks in risk management. We presented the estimated downgrade probability at horizons 1 and 2 of a firm currently rated A for different numbers of time periods and different values of the correlation parameter. Our results confirmed the fact the downgrade probability increases as the horizon increases.

Conclusion

This thesis is an empirical investigation of various estimation methods for the analysis of the dynamics of credit rating matrices. More specifically, the thesis discussed the statistical estimation of the latent factor ordered-Probit model, which is also known as the stochastic credit migration model, a homogeneous nonlinear dynamic panel model with a common unobserved factor, to determine the dynamics and the forecast of credit ratings transition probabilities. Due to computational complexity of the full information maximum likelihood in the latent factor ordered-Probit model, this thesis presents three maximum likelihood estimation methods of the latent factor ordered-Probit model. The first two methods are the two-step estimation and the joint optimization. These methods rely on analytical approximation of the true log-likelihood function of the latent factor ordered-Probit model based on the granularity theory. The third method is maximum composite likelihood estimation which is a new approach to estimate the latent factor ordered-Probit model.

To provide context for the analysis that followed, Chapter one presented the literature review on the dynamics, estimation and modelling of credit rating transition matrices. We explained the general framework of the stochastic migration model and we discussed the statistical inference of the latent factor ordered-Probit model. We reviewed the literature on the granularity-based estimation method and the maximum composite likelihood estimation approach.

Chapter two examined the two estimation methods for the stochastic migration model belong to the class of analytical approximations that are derived from the so-called granularity principle or granularity theory. These methods rely on the Asymptotic Risk Factor (ARF) model under the assumption of an extremely large portfolio. The latent factor ordered-Probit model has been estimated on the French dataset of the Credit Agricole S.A. Bank by these two methods. We assessed the latent factor ordered-Probit model on its ability to link the transition probabilities to an unobserved dynamic risk factor. Our findings shows that the factor ordered-Probit model fits well with the dataset. The implementation of the two-step efficient estimation methodology is easy and less time consuming. The estimates of the latent factor ordered-Probit model parameters obtained by the approximated linear state-space model have been improved by the two-step estimator. The unobserved factor has been recovered by the two-step estimation method and it seems to be close to the dynamics of several leading macro indicators, such as the output gap and the total industrial production variables from the national account group and the short-term interest rate from the financial markets group. The factor interpretation is carried out by using a stress testing analysis. We examined the effects of one-time shocks to the factor on credit migrations and default probabilities. The shocks are used to evaluate stressed migration probabilities and default probabilities during and after crisis of 2008. Our analysis reveals that the shock effects help determine without ambiguity which macroeconomic variable has the same characteristics as the factor and can be interpreted as such. This finding has implications for policy making during economic downturns. The stress-testing results allowed us to interpret the factor, which points to the short-term interest rate in France. We observed a noticeable decrease in the total default rates of all rating categories in the year we introduced a one-time positive shock to the factor and increase in the total default rates of all rating categories when we introduced a one-time negative shock to the factor. The stress scenarios are applied to the economy during the downturn of 2008 and after the recovery in 2010. We observed that the shock effects differ and are more apparent in the post-crisis economy. Therefore, we showed that the model can be used for macro-stress testing purposes.

Chapter three introduced the maximum composite likelihood estimation method for the latent factor ordered-Probit model for credit rating migration matrices. This method provides consistent, and asymptotically normal estimators of the transition probabilities that can be used in forecasting and stress-testing. The main purpose of the composite likelihood methodology is to reduce the computational complexity in models with large datasets and complex dependencies, in a setting where the full likelihood is difficult to construct. This method disregards some of the complex dependencies between observations in the full joint model. For instance, the joint densities can be replaced by a product of pairwise joint densities or a product of conditional densities. Therefore, in the composite likelihood method, the objective function is derived by multiplying a collection of component likelihoods. These methods offer fast and reliable estimation that provides the output quickly. We discussed the identification constraints, order and rank conditions for each of these conditional composite likelihoods. More specifically, we introduced three conditional maximum likelihood functions for the migration model with unobserved AR(1) factor model: the conditional composite likelihood at lag 1, $L_{cc}(\theta)$, the conditional composite likelihood at lag 2, $L_{cc,2}(\theta, \rho)$, and the conditional composite likelihood up to lag 2, which is the combination of the first two functions. We provided the introductory results on the consistency of the estimators and their asymptotic normality. We showed that the asymptotic properties of the maximum composite likelihood estimators with respect to the cross-sectional dimension n and T tend to infinity can be achieved. We undertook a Monte Carlo experiment to assess the small sample performance of the proposed estimators of $L_{cc}(\theta)$. We provided the mean squared error (MSE) and the mean absolute bias of the estimates to give some insight into the accuracy of the estimated parameters. Our results showed that for each value of the correlation parameter ρ, the maximum composite likelihood at lag 1 estimates, have a small estimated bias and MSE, and it decreases as the number of time periods increases. These results are consistent with the asymptotic results on the $\sqrt{T}$consistency of the maximum composite likelihood at lag 1 estimates. We illustrated the finite sample performance of the inferences based on t statistics by plotting the empirical probability distribution function (PDF) of each computed t statistic when testing that each of the parameters equals their true value. Our findings reveal that when T varies the distribution of t statistics is centered on zero and is symmetric, for correlation parameter $\rho = 0.0$ and $\rho = 0.2$. However, when the correlation parameter becomes higher $\rho = 0.4$ and $\rho = 0.7$, for some of the estimated parameters (for e.g.,the intercept), the realized t statistics are shifted from the zero mean due to the smaller variance, suggesting an under-estimation of the true variance. Moreover, we illustrated the ability of the maximum composite likelihood at lag 1 estimation to reproduce the pattern of useful parameters for downgrade risks in risk management. We presented the estimated downgrade probability at horizons 1 and 2 of a firm currently rated A for different numbers of time periods and different values of the correlation parameter. Our results confirmed the fact that the downgrade probability increases as the horizon increases.

References

Acharya, V.V., D. S. and R.K. Sundaram (2002). Pricing Credit Derivatives with Rating Transitions. *Financial Analysts Journal*, 58(3).

Angelini, E., di Tollo, G., and Roli, A. (2008). A neural network approach for credit risk evaluation. *The quarterly review of economics and finance*, 48(4):733–755.

Bae, T. and Kulperger, R. (2009). A Robust Prediction of Corporate Credit Migration. *Available at SSRN 1493362*.

Bangia, A., Diebold, F. X., Kronimus, A., Schagen, C., and Schuermann, T. (2002). Ratings Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing. *Journal of Banking & Finance*, 26(2):445–474.

Basel Committee on Banking Supervision (2000). *Principles for the management of credit risk*. Bank for International Settlements.

Basel Committee on Banking Supervision (2001). The new basel capital accord. Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2003). The new basel capital accord. Technical report, Bank for International Settlements.

Basel Committee on Banking Supervision (2009). Principles for sound stress testing practices and supervision. Technical report.

Belkin, B., Suchower, S., and Forest Jr, L. (1998). A one-parameter representation of credit risk and transition matrices. *CreditMetrics monitor*, 1(3):46–56.

Berteloot, K., Verbeke, W., Castermans, G., Van Gestel, T., Martens, D., and Baesens, B. (2013). A Novel Credit Rating Migration Modelling Approach Using Macroeconomic Indicators. *Journal of Forecasting*, 32(7):654–672.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 192–236.

Besag, J. (1975). Statistical Analysis of Non-lattice Data. *The Statistician*, pages 179–195.

Bharath, S. T. and Shumway, T. (2004). Forecasting default with the KMV-Merton model. In *AFA 2006 Boston Meetings Paper*.

Bhat, C. R., Varin, C., and Ferdous, N. (2010). A Comparison of the Maximum Simulated Likelihood and Composite Marginal Likelihood Estimation Approaches in the Context of the Multivariate Ordered-Response Model. In *Maximum Simulated Likelihood Methods and Applications*, pages 65–106. Emerald Group Publishing Limited.

Bhatia, M., Finger, C. C., and Gupton, G. M. (1997). Credit Metrics–Technical Document. *New York: JP Morgan & Co.*

Black, F. and M.Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654.

Blaschke, M. W. (2001). *Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences*. International Monetary Fund.

Bliss, C. I. (1934). The method of probits. *Science*, 79(2037):38–39.

Brady, B., Vazza, D., and Bos, R. J. (2002). Corporate Defaults Peak in 2002 Amid Record Amounts of Defaults and Declining Credit Quality. *Ratings Performance*.

Bunn, P., Cunningham, A., and Drehmann, M. (2005). Stress testing as a tool for assessing systemic risks. *Bank of England Financial Stability Review*, 18:116–26.

Chandler, R. E. and Bate, S. (2007). Inference for Clustered Data Using the Independence Loglikelihood. *Biometrika*, pages 167–183.

Chauvet, M. and Potter, S. (2005). Forecasting Recessions Using the Yield Curve. *Journal of Forecasting*, 24(2):77–103.

Chava, S. and Jarrow, R. (2002). Default Event Risk, Conditional Diversification and Expected Returns on Defaultable Bonds: A Comparison of Historical and Bond-Implied Estimates. Technical report, Working Paper, Cornell University.

Cheung, S. (1996). Provincial Credit Rating in Canada: An Ordered Probit Analysis. Technical report, Bank of Canada.

Commenges, D. (2002). Inference for multi-state models from interval-censored data. *Statistical methods in medical research*, 11(2):167–182.

Cox, D. R. (1975). Partial Likelihood. *Biometrika*, pages 269–276.

Cox, D. R. and Lewis, P. A. W. (1972). Multivariate point processes. In *Proc. 6th Berkeley Symp. Math. Statist. Prob*, volume 3, pages 401–448.

Cox, D. R. and Reid, N. (2004). A note on Pseudolikelihood Constructed from Marginal Densities. *Biometrika*, 91(3):729–737.

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics*, 28(5):777–795.

Creal, D., Schwaab, B., Koopman, S. J., and Lucas, A. (2014). Observation-Driven Mixed-Measurement Dynamic Factor Models with an Application to Credit Risk. *Review of Economics and Statistics*, 96(5):898–915.

Crouhy, M., Galai, D., and Mark, R. (2000). A comparative analysis of current credit risk models. *Journal of Banking & Finance*, 24(1):59–117.

Davis, R. A. and Yau, C. (2011). Comments on Pairwise Likelihood in Time Series Models. *Statistica Sinica*, pages 255–277.

Dimitras, A. I., Zanakis, S. H., and Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, 90(3):487–513.

Drehmann, M. (2008). Stress tests: Objectives, challenges and modelling choices. *Sveriges Riksbank Economic Review*, 2:60.

Duffie, D. and Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica*, 69(3):633–664.

Duffie, D., Saita, L., and Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, 83(3):635–665.

Duffie, D. and Singleton, K. (1999). Modelling Term Structures of Defaultable Bonds. *Review of Financial Studies*, 12(4):687–720.

Duffie, D. and Wang, K. (2004). *Multi-period Corporate Failure Prediction with Stochastic CCovariates*. National bureau of economic research.

Ederington, L. H. (1985). Classification models and bond ratings. *Financial review*, 20(4):237–262.

Feng, D., Gourieroux, C., and Jasiak, J. (2008). The Ordered Qualitative Model for Credit Rating Transitions. *Journal of Empirical Finance*, 15(1):111–130.

Fermanian, J. D. and Salanie, B. (2004). A Nonparametric Simulated Maximum Likelihood Estimation Method. *Econometric Theory*, 20(04):701–734.

Figlewski, S., Frydman, H., and Liang, W. (2012). Modelling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions. *International Review of Economics & Finance*, 21(1):87–105.

Foulcher, S., Gouriéroux, C., and Tiomo, A. (2005). Latent variable approach to modelling dependence of credit risks: application to French firms and implications for regulatory capital.

Frey, R. and McNeil, A. J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6:59–92.

Frye, J. (2000). Depressing Recoveries. *Risk-London-Risk Magazine Limited*, 13(11):108–111.

Fuh, C.-D. et al. (2006). Efficient likelihood estimation in state space models. *The Annals of Statistics*, 34(4):2026–2068.

Fujii, Y. and Yanagimoto, T. (2005). Pairwise Conditional Score Functions: A Generalization of the Mantel–Haenszel Estimator. *Journal of Statistical Planning and Inference*, 128(1):1–12.

Gagliardini, P. and Gouriéroux, C. (2005). Stochastic Migration Models with Application to Corporate Risk. *Journal of Financial Econometrics*, 3(2):188–226.

Gagliardini, P. and Gouriéroux, C. (2010). Granularity Adjustment for Risk Measures: Systematic versus Unsystematic Risks. In *Proceedings of credit_2010 Conference*.

Gagliardini, P. and Gouriéroux, C. (2013). Granularity Adjustment for Risk Measures: Systematic vs Unsystematic Risks. *Int. J. Approx. Reasoning*, 54(6):717–747.

Gagliardini, P. and Gourieroux, C. (2014a). Efficiency in large dynamic panel models with common factors. *Econometric Theory*, 30(05):961–1020.

Gagliardini, P. and Gourieroux, C. (2014b). *Granularity Theory with Applications to Finance and Insurance*. Cambridge University Press.

Gagliardini, P., Gouriéroux, C., and Monfort, A. (2012). Microinformation, Nonlinear Filtering, and Granularity. *Journal of Financial Econometrics*, 10(1):1–53.

Geweke, J., Marshall, R. C., and Zarkin, G. A. (1986). Mobility Indices in Continuous Time Markov Chains. *Econometrica: Journal of the Econometric Society*, pages 1407–1423.

Glasbey, C. A. (2001). Non-linear Autoregressive Time Series with Multivariate Gaussian Mixtures as Marginal Distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 50(2):143–154.

Godambe, V. P. (1960). An Optimum Property of Regular Maximum Likelihood Estimation. *The Annals of Mathematical Statistics*, 31(4):1208–1211.

Gordy, M. (2003). A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules. *Journal of Financial Intermediation*, 12(3):199–232.

Gordy, M. (2004). Granularity Adjustment in Portfolio Credit Risk Measurement. *Risk Measures for the 21st Century. John Wiley & Sons.*

Gordy, M. and Heitfield, E. (2001). Of Moody's and Merton: A Structural Model of Bond Rating Transitions. *Bank for International Settlements.*

Gordy, M. and Heitfield, E. (2002). Estimating Default Correlations from Short Panels of Credit Rating Performance Data. *Working Paper.*

Gordy, M. B. and Lütkebohmert, E. (2007). Granularity Adjustment for Basel II.

Gouriéroux, C. and Jasiak, J. (2012). Granularity adjustment for default risk factor model with cohorts. *Journal of Banking & Finance*, 36(5):1464–1477.

Gourieroux, C. and Monfort, A. (1995). *Statistics and Econometric Models*, volume 1. Cambridge University Press.

Gourieroux, C. and Monfort, A. (2018). Composite indirect inference with application to corporate risks. *Econometrics and Statistics*, 7:30–45.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrica: Journal of the Econometric Society*, pages 681–700.

Gourieroux, M., Gourieroux, C., Monfort, A., Monfort, D. A., et al. (1996). *Simulation-based econometric methods*. Oxford university press.

Gupton, G. M., Finger, C. C., and Bhatia, M. (1997). *Creditmetrics: Technical document*. JP Morgan & Co.

Hajivassiliou, V. A. and Ruud, P. (1994). Classical Estimation Methods for LDV Models Using Simulation. *Handbook of Econometrics*, 4:2383–2441.

Hamerle, A. and Rösch, D. (2005). Misspecified copulas in credit risk models: how good is Gaussian?

Hamerle, A. and Rösch, D. (2006). Parameterizing credit risk models.

Hanfelt, J. J. (2004a). Composite conditional likelihood for sparse clustered data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(1):259–273.

Hanfelt, J. J. (2004b). Composite conditional likelihood for sparse clustered data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(1):259–273.

Harden, S. J. (2013). *Weighted Composite Likelihoods*. PhD thesis, UCL (University College London).

Hens, N., Aerts, M., Molenberghs, G., Thijs, H., and Verbeke, G. (2005a). Kernel weighted influence measures. *Computational Statistics & Data Analysis*, 48(3):467–487.

Hens, N., Aerts, M., Molenberghs, G., Thijs, H., and Verbeke, G. (2005b). Kernel weighted influence measures. *Computational statistics & data analysis*, 48(3):467–487.

Hjort, N. L., Omre, H., Frisén, M., Godtliebsen, F., Helgeland, J., Møller, J., Jensen, E. B. V., Rudemo, M., and Stryhn, H. (1994). Topics in Spatial Statistics [with Discussion, Comments and Rejoinder]. *Scandinavian Journal of Statistics*, pages 289–357.

Hjort, N. L. and Varin, C. (2008). ML, PL, QL in Markov Chain Models. *Scandinavian Journal of Statistics*, 35(1):64–82.

Israel, R. B., Rosenthal, J. S., and Wei, J. Z. (2001). Finding Generators for Markov Chains via Empirical Transition Matrices, with Applications to Credit Ratings. *Mathematical finance*, 11(2):245–265.

Jafry, Y. and Schuermann, T. (2004). Measurement, Estimation and Comparison of Credit Migration Matrices. *Journal of Banking & Finance*, 28(11):2603–2639.

Jarrow, R. and Turnbull, S. (1992). Credit Risk: Drawing the Analogy. *Risk Magazine*, 5(9):63–70.

Jarrow, R. A., Lando, D., and Turnbull, S. (1997). A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies*, 10(2):481–523.

Jarrow, R. A. and Turnbull, S. (1995). Pricing Derivatives on Financial Securities Subject to Credit Risk. *The Journal of Finance*, 50(1):53–85.

Joe, H. and Lee, Y. (2009). On Weighting of Bivariate Margins in Pairwise Likelihood. *Journal of Multivariate Analysis*, 100(4):670–685.

Kay, R. (1986). A Markov model for analysing cancer markers and disease states in survival studies. *Biometrics*, pages 855–865.

Kent, J. T. (1982). Robust Properties of Likelihood Ratio Tests. *Biometrika*, 69(1):19–27.

Khashman, A. (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Systems with Applications*, 37(9):6233–6239.

Kijima, M. and Komoribayashi, K. (1998). A Markov Chain Model for Valuing Credit Risk Derivatives. *The Journal of Derivatives*, 6(1):97–108.

Kim, J. (1999). A Way to Condition the Transition Matrix on Wind. *Credit Risk Supplement*, pages 1–12.

Koopman, S. J., Kräussl, R., Lucas, A., and Monteiro, A. (2009). Credit Cycles and Macro Fundamentals. *Journal of Empirical Finance*, 16(1):42–54.

Koopman, S. J., Lucas, A., and Monteiro, A. (2008). The Multi-State Latent Factor Intensity Model for Credit Rating Transitions. *Journal of Econometrics*, 142(1):399–424.

Koopman, S. J., Lucas, A., and Schwaab, B. (2012). Dynamic Factor Models with Macro, Frailty, and Industry Effects for US Default Counts: the Credit Crisis of 2008. *Journal of Business & Economic Statistics*, 30(4):521–532.

Krüger, U., Stötzel, M., Trück, S., and Others (2005). Time series Properties of a Rating System Based on Financial Ratios. Technical report, Discussion Paper, Series 2: Banking and Financial Supervision.

Kullback, S. and Leibler, R. (1951). On Information and Sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives research*, 2(2-3):99–120.

Lando, D. (2000). Some elements of rating-based credit risk modelling. *Advanced Fixed-Income Valuation Tools*, pages 193–215.

Lando, D. and Skødeberg, T. M. (2002). Analyzing Rating Transitions and Rating Drift with Continuous Observations. *Journal of Banking & finance*, 26(2):423–444.

Le Cessie, S. and Van Houwelingen, J. C. (1994). Logistic Regression for Correlated Binary Data. *Applied Statistics*, pages 95–108.

Lee, L. F. (1997). Simulation Estimation of Dynamic Switching Regression and Dynamic Disequilibrium Models—Some Monte Carlo Results. *Journal of Econometrics*, 78(1):179–204.

Lele, S. and Taper, M. (2002). A Composite Likelihood Approach to (Co) Variance Components Estimation. *Journal of Statistical Planning and Inference*, 103(1):117–135.

Liang, K.-Y. (1987). Extended mantel-haenszel estimating procedure for multivariate logistic regression models. *Biometrics*, pages 289–299.

Lindsay, B. G. (1988). Composite Likelihood Methods. *Contemporary Mathematics*, 80(1):221–239.

Lindsay, B. G., Yi, G. Y., and J.Sun (2011). Issues and Strategies in the Selection of Composite Likelihoods. *Statistica Sinica*, pages 71–105.

Mählmann, T. (2006). Estimation of Rating Class Transition Probabilities with Incomplete Data. *Journal of Banking & Finance*, 30(11):3235–3256.

Mardia, K. V., Hughes, G., and Taylor, C. (2007). Efficiency of the Pseudolikelihood for Multivariate Normal and Von Mises Distributions. *Preprint*.

Mardia, K. V., Hughes, G., Taylor, C. C., and Singh, H. (2008). A Multivariate Von Mises Distribution with Applications to Bioinformatics. *Canadian Journal of Statistics*, 36(1):99–109.

Martin, R. and Wilde, T. (2002). Credit Portfolio Measurement Unsystematic Credit Risk. *Risk-London-Risk Magazine Limited*, 15(11):123–128.

McNeil, A. J. and Wendin, J. (2007). Bayesian Inference for Generalized Linear Mixed Models of Portfolio Credit Risk. *Journal of Empirical Finance*, 14(2):131–149.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of finance*, 29(2):449–470.

Müller, G. and Czado, C. (2005). An Autoregressive Ordered Probit Model with Application to High-Frequency Financial Data. *Journal of Computational and Graphical Statistics*, 14(2):320–338.

Neyman, J., Scott, E. L., et al. (1948). Consistent Estimates Based on Partially Consistent observations. *Econometrica*, 16(1):1–32.

Ng, C. T., Joe, H., Karlis, D., and Liu, J. (2011). Composite Likelihood for Time Series Models with a Latent Autoregressive Process. *Statistica Sinica*, pages 279–305.

Nickell, P., Perraudin, W., and Varotto, S. (2000). Stability of Rating Transitions. *Journal of Banking & Finance*, 24(1):203–227.

Nickell, P., Perraudin, W., Varotto, S., and Others (2001). *Stability of ratings transitions*. Bank of England.

Pakel, C., Shephard, N., Sheppard, K., and Engle, R. (2017). Fitting vast dimensional time-varying covariance models.

Pinches, G. E. and Mingo, K. A. (1973). A multivariate analysis of industrial bond ratings. *The journal of Finance*, 28(1):1–18.

Rösch, D. (2005). An Empirical Comparison of Default Risk Forecasts from Alternative Credit Rating Philosophies. *International Journal of Forecasting*, 21(1):37–51.

Schönbucher, P. J. (2000). Factor models for portofolio credit risk. Technical report, Bonn Econ Discussion Papers.

Schönbucher, P. J. (2001). Factor Models: Portfolio Credit Risks when Defaults are Correlated. *The Journal of Risk Finance*, 3(1):45–56.

Schönbucher, P. J. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. John Wiley & Sons.

Sorge, M. (2004). Stress-testing financial systems: An overview of current methodologies.

Stefanescu, C., Tunaru, R., and Turnbull, S. (2009). The Credit Rating Process and Estimation of Transition Probabilities: A Bayesian Approach. *Journal of Empirical Finance*, 16(2):216–234.

Stein, M. L., Chi, Z., and Welty, L. (2004). Approximating Likelihoods for Large Spatial Data Sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2):275–296.

Treacy, W. F. and Carey, M. S. (1998). Credit risk rating at large US banks. *Fed. Res. Bull.*, 84:897.

Trück, S. (2008). Forecasting Credit Migration Matrices with Business Cycle Effects—A Model Comparison. *The European Journal of Finance*, 14(5):359–379.

Tuzcuoglu, K. (2017). Composite likelihood estimation of ar-probit model: Application to credit ratings.

Varin, C. (2008). On Composite Marginal Likelihoods. *AStA Advances in Statistical Analysis*, 92(1):1–28.

Varin, C. and Czado, C. (2009). Supplementary Material of the Article: A mixed autoregressive Probit Model for Ordinal Longitudinal Data.

Varin, C., Reid, N., and Firth, D. (2011). An Overview of Composite Likelihood Methods. *Statistica Sinica*, pages 5–42.

Varin, C. and Vidoni, P. (2006). Pairwise Likelihood Inference for Ordinal Categorical Time Series. *Computational Statistics & Data Analysis*, 51(4):2365–2373.

Vasdekis, V. G. S., Cagnone, S., and Moustaki, I. (2012). A Composite Likelihood Inference in Latent Variable Models for Ordinal Longitudinal Responses. *Psychometrika*, 77(3):425–441.

Vecchia, A. V. (1988). Estimation and Model Identification for Continuous Spatial Processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 297–312.

Wang, M. and Williamson (2005). Generalization of the Mantel–Haenszel Estimating Function for Sparse Clustered Binary Data. *Biometrics*, 61(4):973–981.

Wei, J. Z. (2003). A multi-factor, credit migration model for sovereign and corporate debts. *Journal of International Money and Finance*, 22(5):709–735.

Wilde, T. (2001). Basel II Probing Granularity. *Risk-London-Risk Magazine Limited*, 14(8):103–107.

Zhao, Y. and Joe, H. (2005). Composite likelihood estimation in multivariate data analysis. *Canadian Journal of Statistics*, 33(3):335–356.

Appendices

Chapter 2 Appendix

Appendix A

The original adjusted migration probabilities for the period of 2007-2014 are reported in Appendix A.1 and the adjusted migration probabilities for the period of 2007-2014 are reported in A.2.

Appendix A.1

The original adjusted migration probabilities for the period of 2007-2014 are reported in Table A.1.1-A.1.8.

Table A.1.1: The Original Adjusted Transition Matrix for Year 2007 (%)

2007	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	91	53.54	13.13	13.13	10.10	3.03	2.02	4.04	0.00	0.00	0.00	1.01	0.00	0.00	0.00
2	322	2.80	27.33	24.53	11.80	7.76	9.63	3.42	5.59	3.42	1.24	0.31	1.24	0.62	0.31
3	1132	0.88	2.39	28.62	11.93	13.34	16.43	9.36	6.45	4.51	3.36	1.41	0.62	0.62	0.09
4	3181	0.03	0.16	2.77	23.45	14.40	19.96	15.72	11.76	5.94	2.92	1.54	0.53	0.50	0.31
5	2904	0.07	0.03	2.07	4.75	34.92	13.50	15.50	12.36	8.54	3.99	2.10	0.93	0.55	0.69
6	8120	0.01	0.10	0.83	1.98	7.46	29.31	17.40	19.47	12.40	5.99	2.46	0.92	0.84	0.83
7	7263	0.04	0.10	0.54	0.81	3.95	6.83	36.09	18.24	18.08	9.14	3.37	1.25	0.77	0.78
8	6252	0.03	0.03	0.22	0.75	2.91	5.52	11.72	42.25	17.93	11.28	3.82	1.60	0.86	1.07
9	14256	0.02	0.04	0.14	0.32	1.36	3.31	9.48	13.73	38.53	20.19	6.76	2.62	1.98	1.52
10	9207	0.01	0.02	0.09	0.24	1.08	1.80	6.10	9.51	15.90	43.04	10.49	4.95	3.65	3.11
11	3356	0.03	0.00	0.12	0.06	0.60	1.40	3.81	5.90	11.86	19.85	34.51	9.71	6.94	5.21
12	2641	0.04	0.00	0.08	0.38	0.27	1.10	3.10	5.26	9.01	16.74	13.86	31.58	11.66	6.93
13	2451	0.08	0.12	0.00	0.16	0.69	1.43	1.55	3.59	6.00	9.63	9.47	9.71	48.92	8.65
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.2: The Original Adjusted Transition Matrix for Year 2008 (%)

2008	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	63	79.00	3.00	2.00	0.00	14.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	155	2.58	53.55	12.26	9.68	5.81	4.52	3.87	3.23	1.29	1.29	0.00	0.65	1.29	0.00
3	727	0.28	2.61	53.37	18.16	9.22	5.09	3.30	3.16	1.93	0.96	0.96	0.14	0.41	0.41
4	1547	0.06	1.16	5.17	47.32	15.06	12.73	6.08	4.27	3.62	2.00	0.65	0.45	0.65	0.78
5	3504	0.06	0.17	1.43	4.14	52.94	14.27	10.45	5.88	4.79	2.60	1.14	0.43	0.66	1.06
6	5920	0.02	0.03	0.44	1.99	6.72	48.87	18.26	10.29	5.84	4.14	1.30	0.52	0.79	0.78
7	9304	0.02	0.05	0.26	0.56	2.69	7.93	52.43	15.94	9.69	5.30	2.16	0.87	0.76	1.34
8	11695	0.00	0.07	0.13	0.37	1.28	3.67	10.69	50.86	17.29	8.85	3.04	1.10	1.18	1.48
9	13990	0.03	0.05	0.05	0.29	0.79	2.00	5.70	14.20	48.08	16.64	5.83	2.43	2.07	1.85
10	12400	0.00	0.05	0.07	0.10	0.48	1.02	2.69	6.56	16.78	49.09	11.23	5.02	3.20	3.70
11	5401	0.06	0.11	0.02	0.15	0.43	0.94	2.26	4.68	9.65	18.76	40.62	8.79	6.92	6.61
12	3127	0.00	0.03	0.03	0.13	0.26	0.54	1.31	3.42	6.81	12.18	14.58	41.64	12.12	6.94
13	3420	0.03	0.15	0.00	0.09	0.23	0.67	1.32	2.69	4.06	6.70	7.49	8.45	58.22	9.91
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.3: The Original Adjusted Transition Matrix for Year 2009 (%)

2009	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	35	68.32	0.00	10.89	2.97	0.00	0.00	8.91	2.97	2.97	2.97	0.00	0.00	0.00	0.00
2	130	0.77	64.62	16.92	8.46	4.62	3.08	1.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	622	0.48	6.59	61.26	8.84	8.84	2.89	4.34	3.38	1.77	0.32	0.48	0.32	0.32	0.16
4	1340	0.00	0.52	10.82	52.40	15.90	7.76	5.22	2.46	1.94	1.64	0.37	0.30	0.52	0.15
5	3310	0.00	0.18	2.02	8.16	52.91	11.66	10.24	6.47	4.26	2.54	0.60	0.18	0.36	0.42
6	5506	0.02	0.13	0.73	2.56	9.04	50.22	16.75	9.10	5.58	3.20	0.96	0.40	0.80	0.53
7	9427	0.00	0.08	0.19	0.66	3.17	8.78	50.51	16.54	10.31	5.42	2.28	0.69	0.91	0.46
8	12233	0.01	0.10	0.12	0.36	1.28	3.26	10.98	50.37	17.46	9.13	3.45	1.33	1.14	1.01
9	14137	0.02	0.09	0.07	0.21	0.95	1.73	5.37	13.32	48.53	17.29	6.47	2.50	2.00	1.45
10	12988	0.00	0.15	0.09	0.18	0.41	0.85	2.66	5.57	15.77	50.81	11.60	5.19	3.84	2.88
11	6204	0.00	0.03	0.03	0.08	0.44	0.55	2.42	3.74	8.61	18.05	41.68	11.19	8.01	5.17
12	3674	0.03	0.11	0.05	0.05	0.19	0.52	1.69	3.29	5.09	11.35	13.94	40.88	15.13	7.68
13	4144	0.02	0.07	0.05	0.00	0.27	0.39	1.54	2.05	3.26	5.91	6.13	7.00	63.39	9.92
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.4: The Original Adjusted Transition Matrix for Year 2010 (%)

2010	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	29	81.44	3.09	3.09	0.00	3.09	3.09	3.09	3.09	0.00	0.00	0.00	0.00	0.00	0.00
2	230	0.00	85.23	6.52	3.91	0.00	2.17	0.87	0.00	0.87	0.43	0.00	0.00	0.00	0.00
3	706	0.28	4.39	62.75	15.16	8.78	4.11	0.99	0.57	1.56	0.71	0.00	0.28	0.28	0.14
4	1329	0.08	0.68	8.13	57.49	11.96	11.89	4.89	1.88	1.35	1.05	0.45	0.08	0.08	0.00
5	3317	0.12	0.24	2.53	7.15	57.56	13.42	9.04	4.13	2.95	1.57	0.45	0.18	0.45	0.21
6	5145	0.04	0.06	0.52	1.92	9.23	56.92	16.56	7.11	3.58	2.37	0.49	0.27	0.58	0.35
7	9262	0.02	0.03	0.35	0.81	3.12	10.01	59.08	13.98	6.61	3.39	0.97	0.51	0.64	0.49
8	12211	0.01	0.03	0.19	0.42	1.41	3.25	12.30	54.88	15.49	7.36	2.30	0.78	1.01	0.57
9	14343	0.01	0.03	0.04	0.25	0.85	1.66	6.14	15.40	51.14	15.68	4.99	1.46	1.38	0.97
10	13687	0.00	0.01	0.02	0.09	0.39	0.97	3.65	7.35	16.05	52.77	10.16	3.84	2.48	2.21
11	6788	0.00	0.00	0.04	0.03	0.31	0.63	2.52	4.18	9.55	20.83	43.64	8.13	6.00	4.14
12	3937	0.00	0.00	0.05	0.05	0.20	0.36	1.80	3.51	5.92	15.77	16.31	39.01	11.53	5.49
13	5049	0.00	0.02	0.08	0.20	0.26	0.59	1.64	2.81	3.90	5.82	6.75	8.69	60.75	8.48
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.5: The Original Adjusted Transition Matrix for Year 2011 (%)

2011	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	28	66.67	3.92	3.92	6.87	3.92	3.92	0.00	6.86	3.92	0.00	0.00	0.00	0.00	0.00
2	172	0.58	37.80	20.36	14.53	14.53	5.81	1.16	1.16	1.16	1.74	0.00	0.58	0.58	0.00
3	740	0.41	4.46	34.32	20.41	13.92	11.22	4.46	4.73	3.11	1.35	0.27	0.68	0.54	0.14
4	1355	0.22	2.73	5.02	33.87	17.86	16.90	9.89	5.39	3.39	2.80	0.74	0.22	0.59	0.37
5	3358	0.03	0.42	1.55	4.41	30.79	18.25	14.18	11.08	9.68	6.05	1.73	0.60	0.86	0.39
6	5313	0.09	0.13	0.87	2.92	8.41	35.01	17.47	12.95	11.37	6.04	2.30	0.96	1.00	0.49
7	10005	0.01	0.05	0.34	0.73	3.27	10.63	32.41	17.27	15.51	11.63	4.41	1.64	1.33	0.76
8	12624	0.00	0.02	0.40	0.42	1.58	5.85	12.24	33.13	19.73	16.06	5.93	2.03	1.75	0.86
9	13915	0.01	0.01	0.07	0.23	0.84	2.54	7.22	14.13	35.34	22.92	9.64	3.35	2.57	1.14
10	13836	0.00	0.01	0.04	0.12	0.43	1.27	3.84	7.60	15.92	41.10	16.11	6.90	3.79	2.86
11	6588	0.00	0.00	0.02	0.12	0.43	0.59	2.34	4.08	9.59	23.35	34.35	12.80	7.62	4.72
12	3475	0.00	0.00	0.12	0.14	0.26	0.60	1.61	4.29	6.47	16.60	21.18	29.13	11.91	7.68
13	4829	0.00	0.02	0.17	0.14	0.50	0.95	2.32	3.42	4.80	9.71	12.47	12.82	43.38	9.30
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.6: The Original Adjusted Transition Matrix for Year 2012 (%)

2012	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	17	52.48	5.94	17.82	11.88	5.94	0.00	5.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	160	0.63	42.50	17.50	14.38	12.50	4.38	1.25	3.75	1.25	0.63	0.00	0.00	1.25	0.00
3	485	1.24	3.71	36.91	23.51	8.87	7.42	6.80	6.19	1.65	2.68	0.62	0.41	0.00	0.00
4	1069	0.19	0.94	8.14	42.19	16.74	14.87	7.48	3.65	2.71	2.34	0.37	0.09	0.28	0.00
5	2390	0.08	0.17	1.55	7.91	30.21	19.12	16.40	8.20	8.12	4.85	1.59	0.71	0.84	0.25
6	5124	0.02	0.10	0.66	2.38	8.86	37.18	20.14	12.74	9.78	5.50	1.64	0.47	0.35	0.18
7	8002	0.02	0.02	0.20	0.64	3.24	12.43	33.78	18.47	15.02	10.98	2.72	1.29	0.77	0.41
8	10484	0.02	0.02	0.05	0.22	0.88	5.53	13.82	32.51	22.53	16.63	4.73	1.60	1.05	0.42
9	13227	0.00	0.01	0.02	0.13	0.42	2.07	6.32	14.54	37.65	25.73	8.13	2.83	1.21	0.95
10	15630	0.00	0.03	0.04	0.06	0.19	0.81	2.44	5.84	17.72	46.05	16.86	5.75	2.44	1.76
11	8732	0.00	0.01	0.05	0.05	0.10	0.42	0.90	2.24	7.27	25.49	39.50	14.03	5.38	4.55
12	4391	0.00	0.00	0.05	0.14	0.07	0.50	0.87	1.43	3.73	15.35	23.43	35.35	10.73	8.36
13	3737	0.05	0.00	0.11	0.21	0.19	0.51	1.36	1.93	3.21	9.34	11.72	14.40	43.06	13.91
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.7: The Original Adjusted Transition Matrix for Year 2013 (%)

2013	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	14	64.65	0.00	7.07	14.14	0.00	7.07	0.00	7.07	0.00	0.00	0.00	0.00	0.00	0.00
2	105	0.00	65.72	12.38	9.52	3.81	5.71	1.90	0.95	0.00	0.00	0.00	0.00	0.00	0.00
3	380	0.00	4.21	59.47	22.63	5.26	5.26	1.58	0.53	0.53	0.53	0.00	0.00	0.00	0.00
4	939	0.11	1.28	11.82	50.05	14.27	11.71	4.58	2.24	1.49	1.60	0.32	0.21	0.00	0.32
5	1891	0.00	0.16	2.17	10.58	47.54	23.37	8.57	3.12	2.06	1.32	0.74	0.05	0.32	0.00
6	4655	0.00	0.02	0.41	2.73	8.29	49.82	20.88	9.26	4.79	2.58	0.45	0.30	0.34	0.13
7	7129	0.00	0.01	0.14	0.65	2.83	16.02	45.38	18.57	9.64	5.08	0.79	0.32	0.35	0.22
8	9255	0.02	0.01	0.05	0.70	0.91	5.52	17.33	40.86	21.03	9.95	2.12	0.64	0.42	0.43
9	13370	0.02	0.01	0.09	0.13	0.56	2.65	6.94	16.84	43.46	21.85	4.97	1.26	0.66	0.58
10	17717	0.01	0.00	0.01	0.09	0.25	0.80	2.18	6.15	19.00	52.31	12.68	3.72	1.52	1.28
11	9774	0.00	0.02	0.02	0.06	0.11	0.33	0.85	2.48	7.27	28.76	41.87	10.93	3.87	3.44
12	5104	0.00	0.02	0.00	0.04	0.12	0.24	0.45	1.31	3.51	17.05	24.28	34.54	10.40	8.05
13	3256	0.00	0.00	0.06	0.03	0.21	0.71	0.86	1.63	2.89	7.56	11.00	14.22	45.95	14.90
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Table A.1.8: The Original Adjusted Transition Matrix for Year 2014 (%)

2014	ISSUER	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	10	70.01	0.00	10.00	0.00	0.00	0.00	0.00	0.00	0.00	20.00	0.00	0.00	0.00	0.00
2	105	2.86	64.76	15.24	5.71	6.67	0.95	2.86	0.00	0.00	0.95	0.00	0.00	0.00	0.00
3	410	0.49	4.88	54.39	25.37	8.05	3.90	0.73	0.49	0.49	0.73	0.49	0.00	0.00	0.00
4	1005	0.50	0.70	9.15	58.21	15.92	9.35	1.99	1.59	1.19	1.00	0.20	0.10	0.00	0.10
5	1891	0.05	0.11	1.64	8.51	57.06	18.14	7.51	2.86	2.33	1.06	0.11	0.16	0.32	0.16
6	5131	0.00	0.00	0.68	1.68	9.14	48.43	22.33	9.71	4.31	2.81	0.39	0.18	0.21	0.14
7	7648	0.01	0.03	0.18	0.42	2.41	14.78	46.21	19.86	9.54	4.96	0.90	0.24	0.29	0.18
8	9595	0.04	0.00	0.09	0.17	0.94	4.68	16.53	42.67	21.92	9.70	2.30	0.44	0.32	0.20
9	13719	0.02	0.01	0.02	0.11	0.36	1.68	6.07	16.91	45.94	22.01	4.42	1.36	0.63	0.47
10	18489	0.00	0.01	0.02	0.05	0.17	0.55	1.87	5.68	18.12	52.69	14.32	3.76	1.59	1.17
11	9212	0.01	0.03	0.01	0.04	0.08	0.23	0.76	2.03	6.16	26.45	45.93	11.44	3.72	3.10
12	4375	0.00	0.00	0.02	0.00	0.02	0.07	0.53	0.98	3.68	14.19	25.39	39.98	9.42	5.71
13	2985	0.00	0.00	0.07	0.10	0.07	0.30	0.80	1.34	2.18	7.10	11.22	12.63	49.21	14.97
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00

Appendix A.2

The aggregated adjusted migration probabilities for the period of 2007-2014 are reported in Table A.2.1-A.2.8.

Table A.2.1: The Aggregated Adjusted Transition Matrix for Year 2007 (%)

2007	Issuers	A+	A	B+	B	C	D	F
A+	91	53.54	13.13	13.13	10.10	9.09	1.01	0.00
A	322	2.80	27.33	24.53	11.80	20.81	12.42	0.31
B+	1132	0.88	2.39	28.62	11.93	39.13	16.97	0.08
B	3181	0.03	0.16	2.77	23.45	50.09	23.19	0.31
C	18287	0.03	0.22	0.91	1.96	52.82	43.40	0.66
D	38163	0.03	0.03	0.13	0.34	11.81	84.68	2.99
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.2: The Aggregated Adjusted Transition Matrix for Year 2008 (%)

2008	Issuers	A+	A	B+	B	C	D	F
A+	63	79.00	3.00	2.00	0.00	16.00	0.00	0.00
A	155	2.58	53.55	12.26	9.68	14.20	7.73	0.00
B+	727	0.28	2.61	53.37	18.16	17.61	7.56	0.41
B	1547	0.06	1.16	5.17	47.32	33.87	11.64	0.78
C	18728	0.03	0.07	0.54	1.68	69.20	27.38	1.11
D	50033	0.02	0.07	0.07	0.22	7.74	88.28	3.61
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.3: The Aggregated Adjusted Transition Matrix for Year 2009 (%)

2009	Issuers	A+	A	B+	B	C	D	F
A+	35	68.32	0.00	10.89	2.97	8.91	8.91	0.00
A	130	0.77	64.62	16.91	8.46	9.24	0.00	0.00
B+	622	0.48	6.59	61.26	8.84	16.07	6.59	0.16
B	1340	0.00	0.52	10.82	52.40	28.88	7.23	0.15
C	18243	0.01	0.11	0.69	2.59	68.79	27.35	0.47
D	53380	0.01	0.10	0.08	0.19	7.38	89.02	3.22
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.4: The Aggregated Adjusted Transition Matrix for Year 2010 (%)

2010	Issuers	A+	A	B+	B	C	D	F
A+	29	81.44	3.09	3.09	0.00	9.28	3.09	0.00
A	230	0.00	85.23	6.52	3.91	3.04	1.30	0.00
B+	706	0.28	4.39	62.75	15.16	13.88	3.40	0.14
B	1329	0.08	0.68	8.13	57.49	28.73	4.89	0.00
C	17724	0.04	0.07	0.81	2.32	76.72	19.64	0.40
D	56015	0.00	0.02	0.07	0.20	7.95	89.19	2.56
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.5: The Aggregated Adjusted Transition Matrix for Year 2011 (%)

2011	Issuers	A+	A	B+	B	C	D	F
A+	28	66.67	3.92	3.92	6.87	7.84	10.78	0.00
A	172	0.58	37.80	20.36	14.53	21.51	5.22	0.00
B+	740	0.41	4.46	34.32	20.41	29.60	10.68	0.12
B	1355	0.22	2.73	5.02	33.87	44.65	13.13	0.37
C	18676	0.04	0.14	0.71	2.01	53.50	42.99	0.61
D	55267	0.00	0.01	0.14	0.22	9.43	87.13	3.06
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.6: The Aggregated Adjusted Transition Matrix for Year 2012 (%)

2012	Issuers	A+	A	B+	B	C	D	F
A+	17	52.48	5.94	17.82	11.88	11.88	0.00	0.00
A	160	0.63	42.50	17.50	14.38	18.13	6.86	0.00
B+	485	1.24	3.71	36.91	23.51	23.08	11.55	0.00
B	1069	0.19	0.94	8.14	42.19	39.09	9.44	0.00
C	15516	0.03	0.07	0.56	2.33	57.48	39.21	0.31
D	56201	0.01	0.02	0.04	0.12	7.27	89.47	3.08
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.7: The Aggregated Adjusted Transition Matrix for Year 2013 (%)

2013	Issuers	A+	A	B+	B	C	D	F
A+	14	64.65	0.00	7.07	14.14	7.07	7.07	0.00
A	105	0.00	65.72	12.38	9.52	11.42	0.95	0.00
B+	380	0.00	4.21	59.47	22.63	12.10	1.59	0.00
B	939	0.11	1.28	11.82	50.05	30.56	5.86	0.32
C	13675	0.00	0.03	0.51	2.73	71.36	25.20	0.16
D	58476	0.01	0.01	0.04	0.17	7.45	89.63	2.70
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table A.2.8: The Aggregated Adjusted Transition Matrix for Year 2014 (%)

2014	Issuers	A+	A	B+	B	C	D	F
A+	10	70.01	0.00	10.00	0.00	0.00	19.98	0.00
A	105	2.86	64.76	15.24	5.71	10.48	0.95	0.00
B+	410	0.49	4.88	54.39	25.37	12.68	2.19	0.00
B	1005	0.50	0.70	9.15	58.21	27.26	4.08	0.10
C	14670	0.01	0.03	0.54	1.90	71.66	25.70	0.15
D	58375	0.01	0.01	0.03	0.08	6.64	91.03	2.20
F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Appendix B

Appendix B provides the statistic description of the macroeconomic variables. Statistic description of the macroeconomic variables are shown in Table B.

Table B: Descriptive statistics macroeconomic variables

Variables		Mean	StdDev	Min	Max
National accounts	GDP	0.724	1.688	-2.873	2.424
	OUTG	-0.232	1.396	-2.137	2.124
	UNEM	8.844	1.079	7.063	10.291
	EMP	64.121	0.378	63.650	64.925
	IPROD	102.277	5.818	95.517	112.560
	CPI	101.871	3.593	96.376	106.245
	PPI	100.102	3.199	95.425	103.608
Financial markets	SINTS	1.480	1.712	0.080	4.848
	LINTS	3.128	0.945	1.666	4.303
	SPL-S	1.460	1.062	-0.399	2.420
	EXCH	100.236	3.366	96.010	104.244
Stock makets	CAC40	-0.837	21.082	-42	22.300
	VCAC4	23.925	7.399	15.139	37.938
Composite index	CLI	98.748	3.291	91.022	101.562

Appendix C

The estimated transition probabilities obtained by the two-step estimation method over the period 2007 to 2014 are given in Appendix C.1 and appendix C.2 provides the estimated transition probabilities obtained by joint optimization estimation method over the period 2007 to 2014.

Appendix C.1

The estimated transition probabilities obtained by the two-step estimation method over the period 2007 to 2014 are given in Tables C.1.1-C.1.8.

Table C.1.1: Estimated Transition Matrix for Year 2007

2007	A+	A	B+	B	C	D	F
A+	0.74753	0.036174	0.030464	0.025777	0.039879	0.052946	0.067227
A	0.11155	0.21599	0.27629	0.2181	0.16016	0.017909	1.29E-05
B+	0.022129	0.10565	0.24211	0.29118	0.29874	0.040178	1.74E-05
B	2.91E-05	0.002426	0.04194	0.20708	0.6276	0.12092	3.89E-06
C	6.73E-09	6.08E-06	0.000749	0.01909	0.45344	0.52648	0.000231
D	1.14E-10	1.15E-07	2.07E-05	0.000945	0.092334	0.87609	0.030611
F	0	0	0	0	0	0	1

Table C.1.2: Estimated Transition Matrix for Year 2008

2008	A+	A	B+	B	C	D	F
A+	0.78177	0.033161	0.027576	0.023062	0.035135	0.045404	0.053897
A	0.13664	0.23636	0.27728	0.20184	0.13474	0.013151	7.43E-06
B+	0.025834	0.1162	0.25293	0.28983	0.28039	0.034812	1.30E-05
B	0.001321	0.034671	0.20966	0.38882	0.35104	0.014483	2.07E-08
C	2.74E-08	1.82E-05	0.001687	0.033274	0.53558	0.42936	8.94E-05
D	1.06E-10	1.08E-07	1.97E-05	0.00091	0.090539	0.87715	0.031384
F	0	0	0	0	0	0	1

Table C.1.3: Estimated Transition Matrix for Year 2009

2009	A+	A	B+	B	C	D	F
A+	0.79228	0.032151	0.026628	0.022185	0.033629	0.04308	0.050051
A	0.14555	0.24262	0.27674	0.19621	0.12701	0.011862	6.19E-06
B+	0.027143	0.11975	0.25631	0.28913	0.27445	0.033202	1.19E-05
B	0.003708	0.067052	0.28898	0.38972	0.24455	0.005985	3.05E-09
C	4.26E-08	2.56E-05	0.00217	0.039379	0.55994	0.39842	6.50E-05
D	1.04E-10	1.06E-07	1.94E-05	0.000899	0.089962	0.87748	0.031638
F	0	0	0	0	0	0	1

Table C.1.4: Estimated Transition Matrix for Year 2010

2010	A+	A	B+	B	C	D	F
A+	0.80127	0.031255	0.025792	0.021417	0.032323	0.041089	0.046849
A	0.15372	0.24796	0.27591	0.19114	0.12043	0.01082	5.27E-06
B+	0.028341	0.12291	0.25924	0.28842	0.26924	0.031836	1.09E-05
B	0.00847	0.11042	0.35339	0.3598	0.16539	0.002534	5.19E-10
C	6.26E-08	3.44E-05	0.002695	0.045452	0.58006	0.37171	4.88E-05
D	1.01E-10	1.04E-07	1.91E-05	0.00089	0.089456	0.87777	0.031863
F	0	0	0	0	0	0	1

Table C.1.5: Estimated Transition Matrix for Year 2011

2011	A+	A	B+	B	C	D	F
A+	0.765	0.034688	0.029029	0.024419	0.037489	0.049104	0.060272
A	0.12365	0.22635	0.27728	0.21019	0.14714	0.015372	9.80E-06
B+	0.023922	0.11086	0.24759	0.29065	0.28955	0.037416	1.51E-05
B	0.000221	0.010337	0.10551	0.31922	0.51797	0.046743	3.21E-07
C	1.37E-08	1.06E-05	0.001132	0.025381	0.49558	0.47776	0.000145
D	1.10E-10	1.11E-07	2.02E-05	0.000927	0.091433	0.87662	0.030995
F	0	0	0	0	0	0	1

Table C.1.6: Estimated Transition Matrix for Year 2012

2012	A+	A	B+	B	C	D	F
A+	0.77548	0.033744	0.028129	0.023577	0.036023	0.046791	0.05625
A	0.1316	0.23261	0.2774	0.20506	0.13937	0.013959	8.26E-06
B+	0.025094	0.11416	0.25092	0.29018	0.28386	0.035782	1.38E-05
B	0.00069	0.022523	0.16608	0.3709	0.41661	0.023203	6.04E-08
C	2.11E-08	1.48E-05	0.001452	0.030074	0.52071	0.44764	0.000108
D	1.07E-10	1.09E-07	1.99E-05	0.000917	0.090878	0.87695	0.031236
F	0	0	0	0	0	0	1

Table C.1.7: Estimated Transition Matrix for Year 2013

2013	A+	A	B+	B	C	D	F
A+	0.79685	0.0317	0.026206	0.021797	0.032968	0.042068	0.048414
A	0.14964	0.24534	0.27636	0.19366	0.12366	0.011326	5.71E-06
B+	0.027742	0.12134	0.2578	0.28878	0.27182	0.032507	1.14E-05
B	0.00568	0.087076	0.32292	0.37793	0.20248	0.003916	1.26E-09
C	5.18E-08	2.97E-05	0.002422	0.042359	0.57026	0.38487	5.63E-05
D	1.03E-10	1.05E-07	1.93E-05	0.000895	0.089706	0.87763	0.031752
F	0	0	0	0	0	0	1

Table C.1.8: Estimated Transition Matrix for Year 2014

2014	A+	A	B+	B	C	D	F
A+	0.79707	0.031677	0.026185	0.021777	0.032935	0.042018	0.048333
A	0.14984	0.24547	0.27634	0.19354	0.1235	0.0113	5.69E-06
B+	0.027772	0.12142	0.25787	0.28876	0.27169	0.032472	1.14E-05
B	0.005799	0.088172	0.32455	0.37716	0.20049	0.003832	1.21E-09
C	5.23E-08	2.99E-05	0.002435	0.042512	0.57077	0.3842	5.59E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000894	0.089693	0.87764	0.031757
F	0	0	0	0	0	0	1

Appendix C.2

The estimated transition probabilities obtained by joint optimization estimation method over the period 2007 to 2014 are given in Tables C.2.1-C.2.8.

Table C.2.1: Estimated Transition Matrix for Year 2007

2007	A+	A	B+	B	C	D	F
A+	0.735315	0.080232	0.05596	0.032689	0.058337	0.032561	0.004907
A	0.065139	0.18022	0.266255	0.199172	0.266865	0.022345	4.16E-06
B+	0.013297	0.08172	0.206603	0.222275	0.427463	0.048636	6.66E-06
B	0.000156	0.00516	0.04532	0.116054	0.644895	0.188385	3.07E-05
C	4.02E-07	8.74E-05	0.00317	0.021467	0.522798	0.452339	0.000139
D	1.6E-09	6.26E-07	4.55E-05	0.000587	0.091081	0.87712	0.031165
F	0	0	0	0	0	0	1

Table C.2.2: Estimated Transition Matrix for Year 2008

2008	A+	A	B+	B	C	D	F
A+	0.617324	0.097664	0.074035	0.046278	0.091218	0.061314	0.012167
A	0.151014	0.266501	0.277469	0.155054	0.143546	0.006416	3.95E-07
B+	0.022716	0.114329	0.243915	0.228042	0.360599	0.030397	2.42E-06
B	0.001811	0.029712	0.141309	0.220115	0.549934	0.057117	1.28E-06
C	1.76E-06	0.000274	0.007406	0.039995	0.613723	0.338559	4.2E-05
D	1.49E-09	5.89E-07	4.33E-05	0.000563	0.089143	0.878231	0.032019
F	0	0	0	0	0	0	1

Table C.2.3: Estimated Transition Matrix for Year 2009

2009	A+	A	B+	B	C	D	F
A+	0.579422	0.10139	0.078779	0.050241	0.102028	0.072511	0.015629
A	0.186885	0.286966	0.269171	0.13783	0.114892	0.004256	1.88E-07
B+	0.026414	0.125236	0.254038	0.227771	0.340306	0.026233	1.78E-06
B	0.003437	0.045781	0.181448	0.243314	0.488954	0.037064	4.58E-07
C	2.69E-06	0.000378	0.009375	0.047328	0.635911	0.306977	2.9E-05
D	1.46E-09	5.79E-07	4.27E-05	0.000556	0.088574	0.878551	0.032276
F	0	0	0	0	0	0	1

Table C.2.4: Estimated Transition Matrix for Year 2010

2010	A+	A	B+	B	C	D	F
A+	0.538415	0.104402	0.083269	0.054244	0.113715	0.085764	0.020192
A	0.230214	0.303933	0.255149	0.119035	0.088971	0.002698	8.39E-08
B+	0.030873	0.137357	0.264097	0.226486	0.318836	0.022349	1.28E-06
B	0.006495	0.069344	0.22646	0.258733	0.416468	0.022499	1.46E-07
C	4.17E-06	0.000527	0.011941	0.056121	0.656403	0.274984	1.94E-05
D	1.42E-09	5.68E-07	4.2E-05	0.000549	0.087973	0.878885	0.03255
F	0	0	0	0	0	0	1

Table C.2.5: Estimated Transition Matrix for Year 2011

2011	A+	A	B+	B	C	D	F
A+	0.706803	0.085267	0.060759	0.036116	0.066108	0.038692	0.006255
A	0.082344	0.203477	0.274841	0.190837	0.231956	0.016542	2.32E-06
B+	0.01532	0.089458	0.216525	0.22472	0.410699	0.043273	5.16E-06
B	0.000306	0.008449	0.063389	0.143022	0.64082	0.144001	1.42E-05
C	5.93E-07	0.000118	0.003974	0.025391	0.54808	0.422333	0.000103
D	1.57E-09	6.16E-07	4.49E-05	0.000581	0.090581	0.877411	0.031382
F	0	0	0	0	0	0	1

Table C.2.6: Estimated Transition Matrix for Year 2012

2012	A+	A	B+	B	C	D	F
A+	0.674035	0.090399	0.065943	0.039952	0.0752	0.046372	0.008098
A	0.104889	0.228587	0.279831	0.179018	0.195958	0.011715	1.2E-06
B+	0.017832	0.098446	0.227171	0.226644	0.391946	0.037958	3.88E-06
B	0.000618	0.014007	0.088439	0.173561	0.619212	0.104158	5.9E-06
C	9E-07	0.000163	0.005057	0.030323	0.574406	0.389976	7.36E-05
D	1.54E-09	6.06E-07	4.43E-05	0.000574	0.090034	0.877724	0.031622
F	0	0	0	0	0	0	1

Table C.2.7: Estimated Transition Matrix for Year 2013

2013	A+	A	B+	B	C	D	F
A+	0.556979	0.10317	0.081322	0.052471	0.108431	0.079615	0.018011
A	0.210022	0.296957	0.262088	0.127521	0.10009	0.003322	1.21E-07
B+	0.028791	0.131825	0.259656	0.227191	0.328499	0.024036	1.48E-06
B	0.004901	0.057824	0.20607	0.253106	0.449773	0.028324	2.46E-07
C	3.43E-06	0.000454	0.010718	0.052025	0.647574	0.289203	2.33E-05
D	1.44E-09	5.73E-07	4.23E-05	0.000553	0.088243	0.878735	0.032426
F	0	0	0	0	0	0	1

Table C.2.8: Estimated Transition Matrix for Year 2014

2014	A+	A	B+	B	C	D	F
A+	0.559473	0.102988	0.08105	0.052228	0.10772	0.078808	0.017733
A	0.207381	0.29593	0.262948	0.128665	0.10166	0.003416	1.27E-07
B+	0.02852	0.131088	0.259045	0.227271	0.329804	0.024271	1.51E-06
B	0.004716	0.056386	0.203323	0.252176	0.4542	0.029198	2.64E-07
C	3.34E-06	0.000445	0.010561	0.051489	0.646331	0.291146	2.39E-05
D	1.44E-09	5.73E-07	4.23E-05	0.000553	0.08828	0.878715	0.03241
F	0	0	0	0	0	0	1

Appendix D

The stressed transition matrices under positive and negative shock to the factor and under "mutatis mutandis" scenario over the period 2008 to 2014 are reported in Appendix D.1 and D.2 respectively. Appendix D.3 and D.4 report the above results over the period 2010 to 2014 respectively.

The stressed transition matrices under positive and negative shock to the factor and under "ceteris paribus" scenario over the period 2008 to 2014 are reported in Appendix D.5 and D.6 respectively. Appendix D.7 and D.8 report the above results over the period 2010 to 2014 respectively.

Appendix D.1

The stressed transition matrices under positive shock to the factor over the period 2008 to 2014, under "mutatis mutandis" scenario are reported in Table D.1.1-D.1.7.

Table D.1.1: Stressed Transition Matrix for Year 2008-Positive Shock

2008	A+	A	B+	B	C	D	F
A+	0.79975	0.031409	0.025935	0.021548	0.032546	0.041427	0.047387
A	0.1523	0.24706	0.27608	0.19202	0.12155	0.010993	5.42E-06
B+	0.028132	0.12236	0.25874	0.28855	0.27014	0.032067	1.11E-05
B	0.007391	0.10191	0.34324	0.36676	0.17774	0.002953	7.08E-10
C	5.86E-08	3.27E-05	0.002597	0.044361	0.57671	0.37625	5.13E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000891	0.089543	0.87772	0.031825
F	0	0	0	0	0	0	1

Table D.1.2: Stressed Transition Matrix for Year 2009-Positive Shock

2009	A+	A	B+	B	C	D	F
A+	0.79937	0.031448	0.025971	0.021581	0.032601	0.041511	0.047521
A	0.15195	0.24683	0.27612	0.19224	0.12182	0.011036	5.46E-06
B+	0.028081	0.12223	0.25862	0.28858	0.27036	0.032125	1.11E-05
B	0.007142	0.099871	0.34065	0.36838	0.18089	0.003066	7.64E-10
C	5.77E-08	3.23E-05	0.002573	0.044094	0.57587	0.37738	5.19E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000892	0.089564	0.87771	0.031815
F	0	0	0	0	0	0	1

Table D.1.3: Stressed Transition Matrix for Year 2010-Positive Shock

2010	A+	A	B+	B	C	D	F
A+	0.80405	0.030972	0.025529	0.021177	0.031916	0.040474	0.045878
A	0.15635	0.24961	0.27559	0.18953	0.11841	0.010509	5.01E-06
B+	0.028726	0.12391	0.26015	0.28817	0.26761	0.031417	1.06E-05
B	0.010811	0.12716	0.37061	0.34536	0.14416	0.001904	2.93E-10
C	7.06E-08	3.77E-05	0.002882	0.047503	0.58611	0.36343	4.46E-05
D	1.01E-10	1.04E-07	1.90E-05	0.000887	0.089297	0.87786	0.031935
F	0	0	0	0	0	0	1

Table D.1.4: Stressed Transition Matrix for Year 2011-Positive Shock

2011	A+	A	B+	B	C	D	F
A+	0.76623	0.034579	0.028925	0.024322	0.037319	0.048834	0.059795
A	0.12455	0.22708	0.27731	0.20961	0.14623	0.015202	9.61E-06
B+	0.024055	0.11124	0.24797	0.29061	0.28889	0.037224	1.49E-05
B	0.000254	0.011366	0.11172	0.32632	0.50705	0.043292	2.66E-07
C	1.44E-08	1.10E-05	0.001165	0.025892	0.49853	0.47426	0.00014
D	1.10E-10	1.11E-07	2.01E-05	0.000926	0.091369	0.87666	0.031023
F	0	0	0	0	0	0	1

Table D.1.5: Stressed Transition Matrix for Year 2012-Positive Shock

2012	A+	A	B+	B	C	D	F
A+	0.77596	0.0337	0.028087	0.023538	0.035956	0.046685	0.056069
A	0.13198	0.23289	0.27739	0.20481	0.13902	0.013897	8.19E-06
B+	0.025149	0.11431	0.25107	0.29015	0.2836	0.035708	1.37E-05
B	0.000725	0.023299	0.16924	0.37267	0.41166	0.022415	5.57E-08
C	2.15E-08	1.50E-05	0.001469	0.030307	0.52185	0.44625	0.000106
D	1.07E-10	1.09E-07	1.99E-05	0.000916	0.090852	0.87696	0.031247
F	0	0	0	0	0	0	1

Table D.1.6: Stressed Transition Matrix for Year 2013-Positive Shock

2013	A+	A	B+	B	C	D	F
A+	0.79703	0.031682	0.026189	0.021781	0.032941	0.042028	0.048349
A	0.1498	0.24545	0.27634	0.19356	0.12353	0.011305	5.69E-06
B+	0.027766	0.1214	0.25786	0.28877	0.27172	0.032479	1.14E-05
B	0.005775	0.087953	0.32422	0.37731	0.20089	0.003849	1.22E-09
C	5.22E-08	2.99E-05	0.002433	0.042482	0.57067	0.38433	5.59E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000894	0.089696	0.87763	0.031756
F	0	0	0	0	0	0	1

Table D.1.7: Stressed Transition Matrix for Year 2014-Positive Shock

2014	A+	A	B+	B	C	D	F
A+	0.79715	0.03167	0.026178	0.021771	0.032924	0.042002	0.048307
A	0.14991	0.24552	0.27633	0.1935	0.12345	0.011291	5.68E-06
B+	0.027782	0.12144	0.2579	0.28876	0.27165	0.032461	1.13E-05
B	0.005838	0.088525	0.32507	0.37691	0.19986	0.003806	1.19E-09
C	5.24E-08	3.00E-05	0.002439	0.042561	0.57093	0.38398	5.57E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000894	0.089689	0.87764	0.031759
F	0	0	0	0	0	0	1

Appendix D.2

The stressed transition matrices under negative shock to the factor over the period 2008 to 2014, under "mutatis mutandis" scenario are reported in Table D.2.1-D.2.7.

Table D.2.1: Stressed Transition Matrix for Year 2008-Negative Shock

2008	A+	A	B+	B	C	D	F
A+	0.76289	0.034873	0.029207	0.024586	0.037781	0.049569	0.061096
A	0.12211	0.22509	0.27721	0.21119	0.14871	0.015666	1.01E-05
B+	0.023695	0.11021	0.24692	0.29074	0.29068	0.037747	1.53E-05
B	0.000175	0.008763	0.095403	0.30655	0.53595	0.053157	4.42E-07
C	1.25E-08	9.90E-06	0.001077	0.024526	0.49049	0.48374	0.000153
D	1.10E-10	1.12E-07	2.02E-05	0.000929	0.091544	0.87656	0.030948
F	0	0	0	0	0	0	1

Table D.2.2: Stressed Transition Matrix for Year 2009-Negative Shock

2009	A+	A	B+	B	C	D	F
A+	0.78504	0.032851	0.027284	0.022791	0.034668	0.04468	0.052687
A	0.13935	0.23831	0.27715	0.20012	0.13232	0.012742	7.02E-06
B+	0.026232	0.11729	0.25398	0.28963	0.27856	0.034309	1.27E-05
B	0.001835	0.042942	0.23391	0.39323	0.31694	0.011144	1.16E-08
C	3.14E-08	2.02E-05	0.001824	0.03507	0.54324	0.41976	8.11E-05
D	1.05E-10	1.08E-07	1.96E-05	0.000907	0.090361	0.87725	0.031462
F	0	0	0	0	0	0	1

Table D.2.3: Stressed Transition Matrix for Year 2010-Negative Shock

2010	A+	A	B+	B	C	D	F
A+	0.79847	0.031537	0.026055	0.021658	0.032732	0.041709	0.047837
A	0.15112	0.2463	0.27621	0.19275	0.12248	0.011139	5.55E-06
B+	0.02796	0.12191	0.25833	0.28865	0.27088	0.032261	1.12E-05
B	0.006587	0.095175	0.33445	0.37201	0.18842	0.003348	9.14E-10
C	5.55E-08	3.13E-05	0.002519	0.043469	0.57388	0.38004	5.34E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000893	0.089615	0.87768	0.031792
F	0	0	0	0	0	0	1

Table D.2.4: Stressed Transition Matrix for Year 2011-Negative Shock

2011	A+	A	B+	B	C	D	F
A+	0.76377	0.034796	0.029133	0.024517	0.03766	0.049376	0.060752
A	0.12275	0.22562	0.27724	0.21078	0.14806	0.015543	9.99E-06
B+	0.023789	0.11048	0.2472	0.2907	0.29021	0.037609	1.52E-05
B	0.000193	0.00939	0.099522	0.31189	0.52859	0.050412	3.87E-07
C	1.30E-08	1.02E-05	0.001099	0.024878	0.49261	0.48126	0.00015
D	1.10E-10	1.12E-07	2.02E-05	0.000929	0.091498	0.87659	0.030968
F	0	0	0	0	0	0	1

Table D.2.5: Stressed Transition Matrix for Year 2012-Negative Shock

2012	A+	A	B+	B	C	D	F
A+	0.775	0.033788	0.028171	0.023616	0.036091	0.046897	0.056432
A	0.13123	0.23232	0.2774	0.2053	0.13973	0.014022	8.32E-06
B+	0.025038	0.114	0.25076	0.2902	0.28413	0.035857	1.38E-05
B	0.000656	0.021769	0.16295	0.36906	0.42155	0.024015	6.54E-08
C	2.07E-08	1.46E-05	0.001436	0.029842	0.51957	0.44903	0.000109
D	1.08E-10	1.10E-07	1.99E-05	0.000917	0.090903	0.87693	0.031225
F	0	0	0	0	0	0	1

Table D.2.6: Stressed Transition Matrix for Year 2013-Negative Shock

2013	A+	A	B+	B	C	D	F
A+	0.79667	0.031718	0.026223	0.021812	0.032994	0.042109	0.048478
A	0.14948	0.24523	0.27638	0.19377	0.1238	0.011347	5.73E-06
B+	0.027718	0.12127	0.25774	0.2888	0.27193	0.032534	1.14E-05
B	0.005586	0.086206	0.32161	0.37853	0.20408	0.003984	1.31E-09
C	5.14E-08	2.95E-05	0.002411	0.042237	0.56986	0.38541	5.66E-05
D	1.03E-10	1.05E-07	1.93E-05	0.000895	0.089716	0.87762	0.031747
F	0	0	0	0	0	0	1

Table D.2.7: Stressed Transition Matrix for Year 2014-Negative Shock

2014	A+	A	B+	B	C	D	F
A+	0.797	0.031684	0.026192	0.021784	0.032945	0.042034	0.048359
A	0.14978	0.24543	0.27635	0.19358	0.12355	0.011308	5.69E-06
B+	0.027763	0.12139	0.25785	0.28877	0.27173	0.032483	1.14E-05
B	0.005761	0.08782	0.32403	0.3774	0.20113	0.003859	1.22E-09
C	5.21E-08	2.99E-05	0.002431	0.042463	0.57061	0.38441	5.60E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000894	0.089698	0.87763	0.031756
F	0	0	0	0	0	0	1

Appendix D.3

The stressed transition matrices under positive shock to the factor over the period 2010 to 2014, under ”mutatis mutandis” scenario are reported in Table D.3.1-D.3.5.

Table D.3.1: Stressed Transition Matrix for Year 2010-Positive Shock

2010	A+	A	B+	B	C	D	F
A+	0.81825	0.029479	0.024154	0.019926	0.029815	0.037333	0.041039
A	0.17063	0.25792	0.27336	0.18094	0.10814	0.009002	3.83E-06
B+	0.030822	0.12925	0.26483	0.28674	0.25907	0.029288	9.25E-06
B	0.034455	0.23596	0.42085	0.2459	0.062462	0.000374	1.26E-11
C	1.31E-07	6.07E-05	0.004069	0.059457	0.61531	0.32107	2.74E-05
D	9.74E-11	1.01E-07	1.86E-05	0.000871	0.088468	0.87833	0.03231
F	0	0	0	0	0	0	1

Table D.3.2: Stressed Transition Matrix for Year 2011-Positive Shock

2011	A+	A	B+	B	C	D	F
A+	0.77262	0.034006	0.028378	0.023809	0.036427	0.047424	0.057339
A	0.12937	0.23089	0.2774	0.20649	0.14149	0.014338	8.66E-06
B+	0.024765	0.11324	0.25	0.29032	0.28543	0.036227	1.41E-05
B	0.000509	0.018332	0.14786	0.35908	0.44583	0.028398	9.67E-08
C	1.87E-08	1.35E-05	0.001357	0.028713	0.51387	0.45593	0.000117
D	1.08E-10	1.10E-07	1.99E-05	0.00092	0.091031	0.87686	0.031169
F	0	0	0	0	0	0	1

Table D.3.3: Stressed Transition Matrix for Year 2012-Positive Shock

2012	A+	A	B+	B	C	D	F
A+	0.77847	0.033469	0.027868	0.023333	0.035602	0.046131	0.055126
A	0.13397	0.23439	0.27736	0.20354	0.13717	0.013572	7.85E-06
B+	0.025442	0.11512	0.25187	0.29002	0.28222	0.03532	1.34E-05
B	0.000942	0.027744	0.18623	0.38088	0.38556	0.018639	3.65E-08
C	2.39E-08	1.63E-05	0.00156	0.031557	0.5278	0.43897	9.86E-05
D	1.07E-10	1.09E-07	1.98E-05	0.000914	0.090717	0.87704	0.031306
F	0	0	0	0	0	0	1

Table D.3.4: Stressed Transition Matrix for Year 2013-Positive Shock

2013	A+	A	B+	B	C	D	F
A+	0.79798	0.031587	0.026101	0.0217	0.032804	0.041819	0.048013
A	0.15067	0.24601	0.27625	0.19303	0.12284	0.011196	5.60E-06
B+	0.027893	0.12174	0.25817	0.28869	0.27117	0.032335	1.13E-05
B	0.006298	0.092653	0.33098	0.37391	0.19265	0.003513	1.01E-09
C	5.43E-08	3.08E-05	0.002489	0.043129	0.57278	0.38151	5.43E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000893	0.089643	0.87766	0.03178
F	0	0	0	0	0	0	1

Table D.3.5: Stressed Transition Matrix for Year 2014-Positive Shock

2014	A+	A	B+	B	C	D	F
A+	0.79753	0.031632	0.026143	0.021739	0.032869	0.041918	0.048173
A	0.15026	0.24574	0.2763	0.19328	0.12317	0.011247	5.64E-06
B+	0.027833	0.12158	0.25802	0.28873	0.27143	0.032404	1.13E-05
B	0.006044	0.090393	0.32778	0.37556	0.19655	0.00367	1.10E-09
C	5.33E-08	3.04E-05	0.002462	0.042819	0.57178	0.38286	5.51E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000894	0.089668	0.87765	0.031769
F	0	0	0	0	0	0	1

Appendix D.4

The stressed transition matrices under negative shock to the factor over the period 2010 to 2014, under "mutatis mutandis" scenario are reported in Table D.4.1-D.4.5.

Table D.4.1: Stressed Transition Matrix for Year 2010-Negative Shock

2010	A+	A	B+	B	C	D	F
A+	0.78337	0.033009	0.027433	0.02293	0.034906	0.045048	0.053301
A	0.13796	0.23731	0.27722	0.201	0.13355	0.012949	7.23E-06
B+	0.026028	0.11673	0.25344	0.28973	0.27949	0.034565	1.29E-05
B	0.001553	0.038547	0.22147	0.39143	0.33424	0.012753	1.56E-08
C	2.93E-08	1.91E-05	0.001753	0.034145	0.53935	0.42465	8.53E-05
D	1.06E-10	1.08E-07	1.96E-05	0.000909	0.090452	0.8772	0.031422
F	0	0	0	0	0	0	1

Table D.4.2: Stressed Transition Matrix for Year 2011-Negative Shock

2011	A+	A	B+	B	C	D	F
A+	0.75724	0.035361	0.029676	0.025029	0.038559	0.050813	0.063322
A	0.11811	0.22174	0.27696	0.2138	0.15291	0.01647	1.11E-05
B+	0.023102	0.1085	0.24514	0.29093	0.29367	0.038636	1.61E-05
B	9.17E-05	0.005555	0.071796	0.27063	0.57836	0.073571	1.02E-06
C	9.97E-09	8.27E-06	0.000942	0.022373	0.47686	0.49964	0.000179
D	1.12E-10	1.13E-07	2.04E-05	0.000935	0.091837	0.87639	0.030822
F	0	0	0	0	0	0	1

Table D.4.3: Stressed Transition Matrix for Year 2012-Negative Shock

2012	A+	A	B+	B	C	D	F
A+	0.77248	0.034019	0.02839	0.023821	0.036446	0.047455	0.057393
A	0.12926	0.23081	0.2774	0.20656	0.1416	0.014356	8.68E-06
B+	0.024749	0.1132	0.24996	0.29033	0.28551	0.036249	1.41E-05
B	0.000501	0.018145	0.14699	0.35845	0.44724	0.028675	9.89E-08
C	1.86E-08	1.35E-05	0.001352	0.028648	0.51353	0.45634	0.000117
D	1.08E-10	1.10E-07	2.00E-05	0.00092	0.091038	0.87686	0.031166
F	0	0	0	0	0	0	1

Table D.4.4: Stressed Transition Matrix for Year 2013-Negative Shock

2013	A+	A	B+	B	C	D	F
A+	0.79571	0.031812	0.026311	0.021893	0.033132	0.042319	0.048817
A	0.14861	0.24467	0.27646	0.1943	0.12449	0.011457	5.83E-06
B+	0.027592	0.12094	0.25743	0.28887	0.27248	0.032679	1.15E-05
B	0.005117	0.081737	0.31468	0.38153	0.21258	0.004361	1.58E-09
C	4.93E-08	2.86E-05	0.002357	0.041601	0.56772	0.38823	5.83E-05
D	1.03E-10	1.05E-07	1.93E-05	0.000896	0.08977	0.87759	0.031723
F	0	0	0	0	0	0	1

Table D.4.5: Stressed Transition Matrix for Year 2014-Negative Shock

2014	A+	A	B+	B	C	D	F
A+	0.79662	0.031722	0.026227	0.021816	0.033001	0.042119	0.048494
A	0.14943	0.24521	0.27638	0.19379	0.12383	0.011352	5.73E-06
B+	0.027712	0.12126	0.25772	0.2888	0.27195	0.032541	1.14E-05
B	0.005563	0.085989	0.32129	0.37868	0.20448	0.004002	1.32E-09
C	5.13E-08	2.95E-05	0.002409	0.042207	0.56976	0.38554	5.67E-05
D	1.03E-10	1.05E-07	1.93E-05	0.000895	0.089719	0.87762	0.031746
F	0	0	0	0	0	0	1

Appendix D.5

The stressed transition matrices under positive shock to the factor over the period 2008 to 2014, under "ceteris paribus" scenario are reported in Table D.5.1-D.5.7.

Table D.5.1: Stressed Transition Matrix for Year 2008-Positive Shock

2008	A+	A	B+	B	C	D	F
A+	0.79975	0.031409	0.025935	0.021548	0.032546	0.041427	0.047387
A	0.1523	0.24706	0.27608	0.19202	0.12155	0.010993	5.42E-06
B+	0.028132	0.12236	0.25874	0.28855	0.27014	0.032067	1.11E-05
B	0.007391	0.10191	0.34324	0.36676	0.17774	0.002953	7.08E-10
C	5.86E-08	3.27E-05	0.002597	0.044361	0.57671	0.37625	5.13E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000891	0.089543	0.87772	0.031825
F	0	0	0	0	0	0	1

Table D.5.2: Stressed Transition Matrix for Year 2009-Positive Shock

2009	A+	A	B+	B	C	D	F
A+	0.78907	0.032464	0.02692	0.022455	0.034091	0.043789	0.051213
A	0.14276	0.24071	0.27695	0.19796	0.12936	0.012248	6.55E-06
B+	0.026734	0.11865	0.25528	0.28936	0.27628	0.033692	1.22E-05
B	0.002726	0.055292	0.26451	0.39358	0.27595	0.007943	5.57E-09
C	3.72E-08	2.30E-05	0.002009	0.037409	0.55258	0.4079	7.17E-05
D	1.04E-10	1.07E-07	1.95E-05	0.000903	0.09014	0.87738	0.03156
F	0	0	0	0	0	0	1

Table D.5.3: Stressed Transition Matrix for Year 2010-Positive Shock

2010	A+	A	B+	B	C	D	F
A+	0.7847	0.032883	0.027314	0.022819	0.034716	0.044754	0.05281
A	0.13907	0.23811	0.27717	0.2003	0.13257	0.012783	7.06E-06
B+	0.026191	0.11718	0.25387	0.28965	0.27874	0.03436	1.27E-05
B	0.001775	0.042026	0.23139	0.39294	0.32041	0.011453	1.23E-08
C	3.10E-08	2.00E-05	0.00181	0.034882	0.54246	0.42075	8.19E-05
D	1.05E-10	1.08E-07	1.96E-05	0.000907	0.090379	0.87724	0.031454
F	0	0	0	0	0	0	1

Table D.5.4: Stressed Transition Matrix for Year 2011-Positive Shock

2011	A+	A	B+	B	C	D	F
A+	0.78294	0.033049	0.027471	0.022965	0.034967	0.045143	0.05346
A	0.1376	0.23706	0.27724	0.20122	0.13387	0.013003	7.28E-06
B+	0.025976	0.11659	0.2533	0.28976	0.27973	0.034631	1.29E-05
B	0.001488	0.037478	0.2183	0.39082	0.33872	0.013197	1.68E-08
C	2.87E-08	1.89E-05	0.001735	0.03391	0.53834	0.42591	8.64E-05
D	1.06E-10	1.08E-07	1.97E-05	0.000909	0.090475	0.87718	0.031412
F	0	0	0	0	0	0	1

Table D.5.5: Stressed Transition Matrix for Year 2012-Positive Shock

2012	A+	A	B+	B	C	D	F
A+	0.78224	0.033116	0.027534	0.023023	0.035068	0.045299	0.053722
A	0.13702	0.23664	0.27726	0.20159	0.13439	0.013092	7.37E-06
B+	0.025891	0.11636	0.25308	0.2898	0.28012	0.034739	1.30E-05
B	0.001385	0.035773	0.21311	0.38967	0.3461	0.013956	1.91E-08
C	2.79E-08	1.84E-05	0.001706	0.033528	0.53668	0.42798	8.82E-05
D	1.06E-10	1.08E-07	1.97E-05	0.00091	0.090514	0.87716	0.031395
F	0	0	0	0	0	0	1

Table D.5.6: Stressed Transition Matrix for Year 2013-Positive Shock

2013	A+	A	B+	B	C	D	F
A+	0.78195	0.033143	0.027559	0.023047	0.035108	0.045362	0.053827
A	0.13679	0.23647	0.27727	0.20174	0.1346	0.013128	7.40E-06
B+	0.025857	0.11626	0.25299	0.28982	0.28028	0.034783	1.30E-05
B	0.001346	0.035108	0.21104	0.38917	0.34907	0.01427	2.00E-08
C	2.76E-08	1.83E-05	0.001695	0.033375	0.53602	0.4288	8.89E-05
D	1.06E-10	1.08E-07	1.97E-05	0.00091	0.090529	0.87715	0.031388
F	0	0	0	0	0	0	1

Table D.5.7: Stressed Transition Matrix for Year 2014-Positive Shock

2014	A+	A	B+	B	C	D	F
A+	0.78184	0.033153	0.02757	0.023056	0.035124	0.045387	0.053869
A	0.1367	0.2364	0.27727	0.2018	0.13468	0.013142	7.42E-06
B+	0.025843	0.11623	0.25295	0.28982	0.28035	0.0348	1.30E-05
B	0.001331	0.034845	0.21021	0.38896	0.35025	0.014397	2.04E-08
C	2.75E-08	1.82E-05	0.00169	0.033315	0.53575	0.42913	8.92E-05
D	1.06E-10	1.08E-07	1.97E-05	0.00091	0.090535	0.87715	0.031386
F	0	0	0	0	0	0	1

Appendix D.6

The stressed transition matrices under negative shock to the factor over the period 2008 to 2014, under "ceteris paribus" scenario are reported in Table D.6.1-D.6.7.

Table D.6.1: Stressed Transition Matrix for Year 2008-Negative Shock

2008	A+	A	B+	B	C	D	F
A+	0.76289	0.034873	0.029207	0.024586	0.037781	0.049569	0.061096
A	0.12211	0.22509	0.27721	0.21119	0.14871	0.015666	1.01E-05
B+	0.023695	0.11021	0.24692	0.29074	0.29068	0.037747	1.53E-05
B	0.000175	0.008763	0.095403	0.30655	0.53595	0.053157	4.42E-07
C	1.25E-08	9.90E-06	0.001077	0.024526	0.49049	0.48374	0.000153
D	1.10E-10	1.12E-07	2.02E-05	0.000929	0.091544	0.87656	0.030948
F	0	0	0	0	0	0	1

Table D.6.2: Stressed Transition Matrix for Year 2009-Negative Shock

2009	A+	A	B+	B	C	D	F
A+	0.77432	0.033851	0.02823	0.023671	0.036187	0.047048	0.056691
A	0.13069	0.23191	0.2774	0.20564	0.14023	0.014112	8.42E-06
B+	0.02496	0.11378	0.25054	0.29024	0.2845	0.035963	1.39E-05
B	0.00061	0.020731	0.15854	0.36635	0.42856	0.025211	7.32E-08
C	2.01E-08	1.43E-05	0.001413	0.029514	0.51794	0.45101	0.000111
D	1.08E-10	1.10E-07	1.99E-05	0.000918	0.09094	0.87691	0.031209
F	0	0	0	0	0	0	1

Table D.6.3: Stressed Transition Matrix for Year 2010-Negative Shock

2010	A+	A	B+	B	C	D	F
A+	0.7788	0.033438	0.027838	0.023306	0.035555	0.046058	0.055001
A	0.13424	0.23459	0.27735	0.20337	0.13692	0.013529	7.81E-06
B+	0.025481	0.11523	0.25198	0.29	0.28203	0.035269	1.34E-05
B	0.000975	0.028386	0.18855	0.38183	0.38207	0.018178	3.45E-08
C	2.42E-08	1.65E-05	0.001572	0.031727	0.52859	0.438	9.76E-05
D	1.07E-10	1.09E-07	1.98E-05	0.000913	0.090699	0.87705	0.031314
F	0	0	0	0	0	0	1

Table D.6.4: Stressed Transition Matrix for Year 2011-Negative Shock

2011	A+	A	B+	B	C	D	F
A+	0.78058	0.033271	0.027681	0.02316	0.035302	0.045665	0.054337
A	0.13567	0.23565	0.27731	0.20245	0.13561	0.013301	7.58E-06
B+	0.025692	0.11581	0.25255	0.2899	0.28104	0.034994	1.32E-05
B	0.001171	0.032035	0.20113	0.38636	0.36343	0.015875	2.54E-08
C	2.60E-08	1.75E-05	0.00164	0.032648	0.53279	0.43281	9.26E-05
D	1.06E-10	1.08E-07	1.97E-05	0.000911	0.090603	0.87711	0.031356
F	0	0	0	0	0	0	1

Table D.6.5: Stressed Transition Matrix for Year 2012-Negative Shock

2012	A+	A	B+	B	C	D	F
A+	0.78129	0.033205	0.027618	0.023101	0.035202	0.045508	0.054072
A	0.13625	0.23607	0.27729	0.20208	0.13508	0.013211	7.49E-06
B+	0.025777	0.11605	0.25277	0.28985	0.28065	0.034885	1.31E-05
B	0.001259	0.033596	0.20624	0.38789	0.35599	0.015027	2.25E-08
C	2.68E-08	1.79E-05	0.001668	0.033023	0.53446	0.43074	9.07E-05
D	1.06E-10	1.08E-07	1.97E-05	0.000911	0.090565	0.87713	0.031373
F	0	0	0	0	0	0	1

Table D.6.6: Stressed Transition Matrix for Year 2013-Negative Shock

2013	A+	A	B+	B	C	D	F
A+	0.78158	0.033178	0.027593	0.023078	0.035161	0.045445	0.053967
A	0.13648	0.23624	0.27728	0.20194	0.13487	0.013175	7.45E-06
B+	0.025811	0.11614	0.25286	0.28984	0.28049	0.034841	1.31E-05
B	0.001296	0.034238	0.20829	0.38846	0.35302	0.014698	2.14E-08
C	2.72E-08	1.80E-05	0.001679	0.033174	0.53513	0.42991	8.99E-05
D	1.06E-10	1.08E-07	1.97E-05	0.00091	0.090549	0.87714	0.031379
F	0	0	0	0	0	0	1

Table D.6.7: Stressed Transition Matrix for Year 2014-Negative Shock

2014	A+	A	B+	B	C	D	F
A+	0.78169	0.033168	0.027583	0.023069	0.035145	0.04542	0.053925
A	0.13657	0.23631	0.27728	0.20188	0.13479	0.013161	7.44E-06
B+	0.025825	0.11618	0.2529	0.28983	0.28043	0.034824	1.31E-05
B	0.001311	0.034497	0.20912	0.38868	0.35183	0.014569	2.10E-08
C	2.73E-08	1.81E-05	0.001684	0.033234	0.5354	0.42958	8.96E-05
D	1.06E-10	1.08E-07	1.97E-05	0.00091	0.090543	0.87714	0.031382
F	0	0	0	0	0	0	1

Appendix D.7

The stressed transition matrices under positive shock to the factor over the period 2010 to 2014, under "ceteris paribus" scenario are reported in Table D.7.1-D.7.5.

Table D.7.1: Stressed Transition Matrix for Year 2010-Positive Shock

2010	A+	A	B+	B	C	D	F
A+	0.81825	0.029479	0.024154	0.019926	0.029815	0.037333	0.041039
A	0.17063	0.25792	0.27336	0.18094	0.10814	0.009002	3.83E-06
B+	0.030822	0.12925	0.26483	0.28674	0.25907	0.029288	9.25E-06
B	0.034455	0.23596	0.42085	0.2459	0.062462	0.000374	1.26E-11
C	1.31E-07	6.07E-05	0.004069	0.059457	0.61531	0.32107	2.74E-05
D	9.74E-11	1.01E-07	1.86E-05	0.000871	0.088468	0.87833	0.03231
F	0	0	0	0	0	0	1

Table D.7.2: Stressed Transition Matrix for Year 2011-Positive Shock

2011	A+	A	B+	B	C	D	F
A+	0.80818	0.030546	0.025136	0.020817	0.03131	0.039561	0.044451
A	0.16035	0.25204	0.27504	0.1871	0.11542	0.010058	4.64E-06
B+	0.029313	0.12543	0.2615	0.28779	0.26516	0.030796	1.02E-05
B	0.015371	0.15494	0.39249	0.32018	0.1158	0.001222	1.22E-10
C	8.44E-08	4.33E-05	0.003184	0.050713	0.59489	0.35113	3.88E-05
D	9.98E-11	1.03E-07	1.89E-05	0.000882	0.08906	0.878	0.032041
F	0	0	0	0	0	0	1

Table D.7.3: Stressed Transition Matrix for Year 2012-Positive Shock

2012	A+	A	B+	B	C	D	F
A+	0.80405	0.030972	0.025529	0.021177	0.031916	0.040474	0.045878
A	0.15635	0.24961	0.27559	0.18953	0.11841	0.010509	5.01E-06
B+	0.028726	0.12391	0.26015	0.28817	0.26761	0.031417	1.06E-05
B	0.010811	0.12716	0.37061	0.34536	0.14416	0.001904	2.93E-10
C	7.06E-08	3.77E-05	0.002882	0.047503	0.58611	0.36343	4.46E-05
D	1.01E-10	1.04E-07	1.90E-05	0.000887	0.089297	0.87786	0.031935
F	0	0	0	0	0	0	1

Table D.7.4: Stressed Transition Matrix for Year 2013-Positive Shock

2013	A+	A	B+	B	C	D	F
A+	0.80239	0.031142	0.025687	0.021321	0.03216	0.040842	0.046459
A	0.15477	0.24862	0.27579	0.1905	0.11962	0.010695	5.16E-06
B+	0.028495	0.12331	0.25961	0.28832	0.26859	0.031668	1.08E-05
B	0.009347	0.11694	0.3605	0.35427	0.15668	0.002262	4.13E-10
C	6.57E-08	3.57E-05	0.002768	0.046264	0.5825	0.36839	4.71E-05
D	1.01E-10	1.04E-07	1.91E-05	0.000889	0.089392	0.87781	0.031892
F	0	0	0	0	0	0	1

Table D.7.5: Stressed Transition Matrix for Year 2014-Positive Shock

2014	A+	A	B+	B	C	D	F
A+	0.80172	0.03121	0.02575	0.021379	0.032258	0.04099	0.046693
A	0.15414	0.24823	0.27586	0.19089	0.12011	0.01077	5.23E-06
B+	0.028402	0.12307	0.25939	0.28838	0.26898	0.031769	1.09E-05
B	0.008812	0.113	0.35627	0.35763	0.16187	0.002422	4.74E-10
C	6.38E-08	3.49E-05	0.002724	0.045775	0.58104	0.37038	4.81E-05
D	1.01E-10	1.04E-07	1.91E-05	0.000889	0.089431	0.87779	0.031875
F	0	0	0	0	0	0	1

Appendix D.8

The stressed transition matrices under negative shock to the factor over the period 2010 to 2014, under "ceteris paribus" scenario are reported in Table D.8.1-D.8.5.

Table D.8.1: Stressed Transition Matrix for Year 2010-Negative Shock

2010	A+	A	B+	B	C	D	F
A+	0.78337	0.033009	0.027433	0.02293	0.034906	0.045048	0.053301
A	0.13796	0.23731	0.27722	0.201	0.13355	0.012949	7.23E-06
B+	0.026028	0.11673	0.25344	0.28973	0.27949	0.034565	1.29E-05
B	0.001553	0.038547	0.22147	0.39143	0.33424	0.012753	1.56E-08
C	2.93E-08	1.91E-05	0.001753	0.034145	0.53935	0.42465	8.53E-05
D	1.06E-10	1.08E-07	1.96E-05	0.000909	0.090452	0.8772	0.031422
F	0	0	0	0	0	0	1

Table D.8.2: Stressed Transition Matrix for Year 2011-Negative Shock

2011	A+	A	B+	B	C	D	F
A+	0.79422	0.03196	0.026449	0.02202	0.033348	0.042649	0.049351
A	0.14728	0.24378	0.27659	0.19513	0.12558	0.011632	5.98E-06
B+	0.027396	0.12042	0.25695	0.28899	0.27334	0.032906	1.17E-05
B	0.004454	0.075093	0.30364	0.38559	0.22621	0.005011	2.10E-09
C	4.63E-08	2.72E-05	0.002274	0.040623	0.56436	0.39266	6.11E-05
D	1.03E-10	1.06E-07	1.93E-05	0.000897	0.089853	0.87754	0.031686
F	0	0	0	0	0	0	1

Table D.8.3: Stressed Transition Matrix for Year 2012-Negative Shock

2012	A+	A	B+	B	C	D	F
A+	0.79847	0.031537	0.026055	0.021658	0.032732	0.041709	0.047837
A	0.15112	0.2463	0.27621	0.19275	0.12248	0.011139	5.55E-06
B+	0.02796	0.12191	0.25833	0.28865	0.27088	0.032261	1.12E-05
B	0.006587	0.095175	0.33445	0.37201	0.18842	0.003348	9.14E-10
C	5.55E-08	3.13E-05	0.002519	0.043469	0.57388	0.38004	5.34E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000893	0.089615	0.87768	0.031792
F	0	0	0	0	0	0	1

Table D.8.4: Stressed Transition Matrix for Year 2013-Negative Shock

2013	A+	A	B+	B	C	D	F
A+	0.80016	0.031368	0.025897	0.021513	0.032486	0.041336	0.047243
A	0.15268	0.2473	0.27603	0.19178	0.12125	0.010947	5.38E-06
B+	0.028188	0.12251	0.25888	0.28851	0.2699	0.032005	1.10E-05
B	0.007667	0.10414	0.346	0.36497	0.17439	0.002835	6.52E-10
C	5.96E-08	3.31E-05	0.002623	0.044651	0.57761	0.37503	5.06E-05
D	1.02E-10	1.05E-07	1.92E-05	0.000891	0.089519	0.87774	0.031835
F	0	0	0	0	0	0	1

Table D.8.5: Stressed Transition Matrix for Year 2014-Negative Shock

2014	A+	A	B+	B	C	D	F
A+	0.80083	0.0313	0.025834	0.021456	0.032388	0.041188	0.047006
A	0.1533	0.2477	0.27596	0.1914	0.12076	0.01087	5.31E-06
B+	0.02828	0.12275	0.2591	0.28845	0.26951	0.031904	1.10E-05
B	0.00814	0.10788	0.35047	0.36191	0.16895	0.002651	5.69E-10
C	6.14E-08	3.39E-05	0.002666	0.04513	0.57908	0.37304	4.95E-05
D	1.02E-10	1.04E-07	1.91E-05	0.00089	0.089481	0.87776	0.031852
F	0	0	0	0	0	0	1

Chapter 3 Appendix

Appendix E

In Appendix E.1., the elements of matrix P which is the expectation of matrix P_t and Appendix E.2. the elements of matrix $P(1)$ which is the expectation of matrix $P_t^{(1)} = P_t P_{t-1}$ are provided.

E.1. Elements of Matrix P

We have:

$$y_{i,t}^* = \beta_l f_t + \delta_l + \sigma_l u_{i,t},$$

where $u_{i,t} \sim N(0,1)$ and $f_t \sim N(0,1)$ are independent. If $y_{i,t-1} = l$, then $y_{i,t}^* \sim N(\delta_l, \sigma_l^2 + \beta_l^2)$. We have that:

$$p_{kl}(\theta) = \Phi\left(\frac{c_{k+1} - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2}}\right) - \Phi\left(\frac{c_k - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2}}\right)$$

E.2. Elements of Matrix $P(1)$

$$
\begin{aligned}
&P(1) = E[P(f_t;\theta).P(f_{t-1};\theta)],\\
&E[P(\rho f_{t-1} + \sqrt{1-\rho^2}\ \eta_t;\theta)].P(f_{t-1};\theta)],
\end{aligned}
$$

Where f_{t-1} and η_t are independent, $\eta_t \sim N(0,1)$ and $f_{t-1} \sim N(0,1)$. Thus:

$$
\begin{aligned}
&= E_{f_{t-1}}\ E_{\eta_t}\Bigg[P(\rho f_{t-1} + \sqrt{1-\rho^2}\ \eta_t;\theta).P(f_{t-1};\theta)|f_{t-1}\Bigg],\\
&= E_{f_{t-1}}\Bigg[E_{\eta_t}\Bigg[P(\rho f_{t-1} + \sqrt{1-\rho^2}\ \eta_t;\theta)|f_{t-1}\Bigg].P(f_{t-1};\theta)\Bigg],\\
&= E_{f_{t-1}}\Bigg[A.B\Bigg],
\end{aligned}
$$

Where matrix A has components as follows:

$$
\begin{aligned}
a_{kl,t} &= \mathbb{P}\Bigg[c_k < y^*_{i,t} < c_{k+1}|\ y_{i,t-1} = l, f_t\Bigg]\\
&= \mathbb{P}\Bigg[c_k < \delta_l + \beta_l f_t + \sigma_l u_{i,t} < c_{k+1}|\ f_t\Bigg]\\
&= \mathbb{P}\Bigg[c_k < \delta_l + \beta_l f_{t-1} + \beta_l\sqrt{1-\rho^2}\ \eta_t + \sigma_l u_{i,t} < c_{k+1}|\ \eta_t, f_{t-1}\Bigg]\\
a_{kl,}(f_{t-1};\theta) &= E_{\eta_t}\Bigg[p^{(1)}_{kl,t}|f_{t-1}\Bigg]\\
&= \Phi\left(\frac{c_{k+1} - \delta_l - \beta_l\rho f_{t-1}}{\sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}}\right) - \Phi\left(\frac{c_k - \delta_l - \beta_l\rho f_{t-1}}{\sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}}\right)
\end{aligned}
$$

Matrix B has elements as follows:

$$
p_{kl,t-1} = p_{kl}(f_{t-1};\theta) = \Phi\left(\frac{c_{k+1} - \delta_l - \beta_l f_{t-1}}{\sigma_l}\right) - \Phi\left(\frac{c_k - \delta_l - \beta_l f_{t-1}}{\sigma_l}\right)
$$

Hence $E_{f_{t-1}}(A.B) = \int AB\phi(f_{t-1})df_t$ where $\phi(f_{t-1})$ is the density of standard normal. Therefore,

$$\begin{aligned}
p_{kl}(1;\theta,\rho) &= \int \sum_{j=1}^{K} a_{k,j}(f;\theta)\ p_{j,l}(f;\theta)\phi(f)df \\
&= \int \sum_{j=1}^{K} \left[\Phi\left(\frac{c_{k+1}-\delta_j-\beta_j\rho f}{\sqrt{\sigma_j^2+\beta_j^2(1-\rho^2)}}\right) - \Phi\left(\frac{c_k-\delta_j-\beta_j\rho f}{\sqrt{\sigma_j^2+\beta_j^2(1-\rho^2)}}\right)\right] \\
&* \left[\Phi\left(\frac{c_{j+1}-\delta_l-\beta_l f}{\sigma_l}\right) - \Phi\left(\frac{c_j-\delta_l-\beta_l f}{\sigma_l}\right)\right] * \phi(f)df
\end{aligned}$$

Appendix F

The proof of Proposition 1 is given in Appendix F.1. and The proof of Proposition 2 is given in Appendix F.2.

F.1. Proof of Proposition 1

We know that the following identifying functions are available:

$$(1)\quad \frac{c_k - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2}} \quad \forall\ k = 1, ..., K-1,$$
$$l = 1, ..., K,$$

therefore parameter ρ is not identifiable. Moreover, parameters σ_l^2 cannot be distinguished from β_l^2. Let us denote their sum by γ_l:

$$\gamma_l = \sqrt{\sigma_l^2 + \beta_l^2}$$

There are $K(K-1)$ identifying functions (1), that we would like to use to identify $(K-1)$ values of c_k, K values of δ_l and K values of γ_l, i.e. $3K-1$ unknowns. We follow Gagliardini and Gouriéroux (2013) and add the identifying constraints:

$$c_1 = 0, \quad \gamma_1 = 1$$

Next, we proceed as follows:

(a) $K = 1$, we identify $\frac{\delta_l}{\gamma_l}$. Given that $\gamma_1 = 1$, we get δ_1 identified.

(b) For $l = 1$, we have $\gamma_1 = 1$, hence we identify $c_k - \delta_1$, given (a). Therefore, all thresholds $c_k, \quad k = 1, ..., K-1$ are identified.

(c) We have:

$$\frac{c_k - \delta_l}{\gamma_l} = \frac{c_k}{\gamma_l} - \frac{\delta_l}{\gamma_l}$$

From (a), we identify $\frac{c_1}{\gamma_l}$

From (b), we identify γ_l, $l = 1, ..., K$.

From (c), we identify δ_l, $l = 1, ..., K$

F.2. Proof of Proposition 2

We have the following identifying functions:

$$(1) \quad \frac{c_k - \delta_l}{\sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}}$$

$$(2) \quad \frac{\epsilon \beta_l \rho}{\sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}}$$

$$(3) \quad \frac{c_k - \delta_l}{\sigma_l}$$

$$(4) \quad \frac{\epsilon \beta_l}{\sigma_l}$$

We define

$$\gamma_l = \sqrt{\sigma_l^2 + \beta_l^2(1-\rho^2)}$$

and has the identifying constraints:

$$\gamma_1 = 1, \quad c_1 = 0,$$

Then we proceed as follows:

(a) For $K = 1$, given $c_1 = 0$

From equation (1) we identify $\frac{\delta_l}{\gamma_l}$

(b) For $K = 1$, given $c_1 = 0$

From equation (3) we identify $\frac{\delta_l}{\sigma_l}$

(c) Given that $\gamma_1 = 1$ and $c_1 = 0$ parameter δ_1 is identified.

(d) It follows that parameter σ_1 is identified.

(e) For $l = 1$, from equation (1) we have

$$\frac{c_k - \delta_1}{\gamma_1} = c_k$$

hence c_k, $k = 1, ..., K-1$ are identified.

(f) From equation (1) We have

$$\frac{c_k}{\gamma_l} - \frac{\delta_l}{\gamma_l},$$

and by (a), ratios $\frac{c_k}{\gamma_l}$ are identified. (g) From (f) and (e), parameters γ_l, $l = 1, ..., K$ are identified.

(h) From (a) and (g), parameters δ_l, $l = 1, ..., K$ are identified.

(i) From (b) and (h), parameters σ_l, $l = 1, ..., K$ are identified.

(j) From equation (4) and (i), parameters $\epsilon\beta_l$, $l = 1, ..., K$ are identified.

(k) We get the ratios $\frac{\epsilon\beta_l\rho}{\gamma_l}$ and given (g) we identify $\epsilon\beta_l\rho$, $l = 1, ..., K$

(l) From (j) and (k), we identify parameter ρ.

Appendix G

In Appendix G, the list of regularity conditions for the asymptotic analysis is provided. In addition to assumptions in the text, the below regularity conditions are given to derive the large sample properties of the estimators. Moreover, these regularity conditions need to be satisfied to provide the validity of the proposition 1, 2 and 3.

G.1. Regularity Conditions

RC.1: The parameter sets $\theta \subset \mathbb{R}^q$, $\alpha \subset \mathbb{R}^q$ and $\Theta \subset \mathbb{R}^p$ are compact. The true parameters θ_0 and α_0 are interior points of sets θ, α and Θ, respectively.

RC.2: The innovations are independent and identically distributed over i and t, that is, $u_{i,t} \overset{i.i.d}{\sim} \mathcal{N}(0,1)$.

RC.3: The conditional composite log-likelihood function

$$L_{cc,n,T}(\alpha) = \sum_{k=1}^{K} \sum_{l=1}^{K} \sum_{t=2}^{T} n_{kl,t} \; log \; p_{kl}(c, \delta, \gamma)$$

is continuous in α.

RC.4: The true but unknown parameter value α_0 is identified.

RC.5: $E_{\alpha_0}[\; log \; p_{kl}(c, \delta, \gamma)]$ exists.

RC.6: The conditional composite log-likelihood function is such that $(\frac{1}{n}) \; L_{cc,n,T}(\alpha)$ converges almost surely to $E_{\alpha_0}[\; log \; p_{kl}(c, \delta, \gamma)]$ uniformly in $\alpha \in \Theta$.

RC.7: The conditional composite log-likelihood function $L_{cc,n,t}(\alpha)$ is twice continuously differentiable in an open neighborhood of α_0.

RC.8: A composite likelihood estimator $\hat{\alpha}_{n,T}$ of α is a unique solution to

$$L_{cc,n,T}(\alpha) = \sum_{k=1}^{K} \sum_{l=1}^{K} \sum_{t=2}^{T} n_{kl,t} \; log \; p_{kl}(c, \delta, \gamma)$$

RC.9: The matrix

$$J_0 = -\sum_{k=1}^{K} \sum_{l=1}^{K} p_{l0} \; p_{kl}(\alpha_0) \; \frac{\partial^2 log \; p_{kl}(\alpha_0)}{\partial \alpha \partial \alpha'}$$

exists and is nonsingular.

Property 1. The assumption RC.1. and RC.3. ensure the existence of a composite maximum estimator $\hat{\alpha}_{n,T}$. It is obtained by maximizing the $L_{cc,n,T}(\alpha)$ or equivalently by maximizing $(\frac{1}{nT}\ L_{cc,n,T}(\alpha))$. Moreover, based on assumption RC.6. then the solution $\hat{\alpha}_{n,T}$ converges to the solution to the limit problem

$$\lim_{n,T\to\infty} \frac{1}{nT} \sum_{k=1}^{K} \sum_{l=1}^{K} \sum_{t=2}^{T} \left[n_{kl,t}\ log\ p_{kl}(c, \delta, \gamma) \right]$$

Now, properties of the Kullback information measure together with the identification constraints on α_0 (see assumption RC.4.) imply that the solution to the limit problem is unique and equal to α_0. Thus, $\hat{\alpha}_{n,T}$ converges almost surly to α_0.