The Royal Institute of Technology (KTH)
School of Engineering Sciences
div. of Mathematical Statistics
Master Thesis in Financial Mathematics

INCORPORATING THE STATE OF THE ECONOMY IN RATINGS MIGRATION FORECASTING

Cecilia Petterson
Supervisor: Ass. Prof. Harald Lang
2008-2009

Credit risk is considered the dominant source of risk in the world of banking and finance and is therefore the attention of both private and public actors. Thus, the need of a method to accurately estimate credit ratings migration probabilities is a necessity, but yet there is none deemed adequate. The aim of this thesis is to propose a proof of concept and suggest a method that, applied on internal rating data, estimates ratings migration probabilities while incorporating the state of the economy. The field of modelling of ratings migration as currently known is presented and aspects of importance will be highlighted. Further, the method of discrete-time Maximum Likelihood estimation is proposed and empirically applied with the incorporation of a macro economic variable. As a result, less diagonal-dominant transition matrices that captures generated movements. Thus, the method renders possible stress testing of transition matrices for historical as well as fictive scenarios.

Keywords: Credit ratings migration, Probability of default, Internal ratings, Maximum Likelihood, State of the economy, Stress testing

INDEX

INTRODUCTION

In the business of banking and finance, credit risk is considered the dominant source of risk and is therefore the attention of many,[i] both private and public actors. Changes in credit quality of counterparties, i.e. borrowers, are known as credit migration and are cardinal input to many risk management applications. As such, there are portfolio risk assessment, pricing of credit derivatives, modelling the term structure of credit risk premium and assessment of regulatory capital.[ii] Thus, in order to measure and monitor credit risk, there are many regulations concerning this constant advancing field. The most important are perhaps the Basel capital accords, which are recommendations on banking laws and regulations issued by the Basel Committee on Banking Supervision to promote stability in the financial system. Due to the fact that the Basel capital accords soon will allow certain financial institutions to use an internal ratings-based (IRB) approach to determine capital requirements for exposures, a method to accurately estimate credit ratings migration probabilities is of great necessity to financial institutions in general and to banks in particular. Empirically, the need has also been indicated by the financial situation that erupted all around the globe during the past year.

The most commonly used methods for estimating credit ratings migration have proven inefficiencies, and the trend in research is to propose more and more complex methods. These, due to their intricate nature, are hard to grasp as well as to implement. In addition, these methods lull one into a false sense of security, when implying that they mirror the world to a greater extent than they actually do. In order to find a suitable method, one has to understand the basics of credit quality dynamics. These are often described as a stochastic process of credit grades, represented by a matrix whose elements represent current and future states of credit grades that are independent of the past states, i.e. a Markov chain. The independence of the process history is debated, but the general idea of a matrix, with elements representing a current rating grade and probabilities for ending up in another grade during a certain period of time, is representative. By intuition, one deems that these transitions are dependent on time, i.e. conditional on the state of the economy. However, there is no clearly discernible way of how to incorporate this. Even though there are many proposed methods, none have had great impact. Incorporating the state of the economy would render possible to stress test transition probabilities, and to see the impact of possible scenarios never historically experienced. This is a great advantage in financial applications, since history is known not to only repeat itself, but to do so due to constantly new reasons.

The aim of this thesis is to give a proof of concept and to propose a method, which applied on internal rating data provided by a bank, estimates ratings migration probabilities while incorporating the state of the economy. In the forthcoming text, an overview of the field of modelling of ratings migration matrices as currently known is presented and aspects of utmost importance to the aim of this thesis will be highlighted. As a starting point, aspects regarding the data the estimations are to be applied upon are discussed, since these often bring about limitations not always regarded in literature. When established what aspects that ought to be taken into consideration when attempting to develop a method superior to the commonly applied – as well as those proposed in recent years – a refined method founded on the facts emerged is proposed; the discrete-time Maximum Likelihood (ML) estimation. The method is empirically applied on internal rating data supplied by Skandinaviska Enskilda Banken AB, publ. (SEB) in two steps; first in the case of time-homogeneity, i.e. there is no time-dependence, and compared with the widely applied cohort method. Second – in an extended model that incorporates the

[i] Typically, the risk taxonomy apart from credit risk includes market and operational risk.
[ii] Jafry & Schuermann (2004) and Mählmann (2006)

state of the economy – in the application of different scenarios that demonstrates the possibilities of stress testing. The thesis concludes with remarks on the result.

THEORETICAL FIELD

When modelling credit ratings migration there are numerous ways to proceed. The first – and in literature often disregarded – step is to consider what data are available to base computations on. This is decisive when choosing if and what Markov model that is adequate for our modelling purposes. By intuition, one often deems that a credit rating may change at any point in time, and that therefore a discrete approach is incorrect – i.e. evaluating and/or reading off the rating at predetermined times. However, one may argue that the opposite – to continuously monitor the transitions – does not adequately describe what is actually taking place, particularly when using internal rating data.[iii] The monitoring of a borrower is in fact not continuous, why the exact time of a rating migration is not observed, and a credit grade could change without anyone noticing for quite a while. It should be noted that the Basel Committee on Banking Supervision, when formulating the minimum requirements for the application of the IRB approach, acknowledges the impossibility of a continuous monitoring of small and medium-sized borrower ratings. Banks are only required to refresh their borrower and facility ratings “at least on an annual basis”[iv], whereas the frequency of reviews must be risk-sensitive.

Concerning external rating data, there are other aspects to consider that will not be discussed further due to the aim of this thesis. Regarding internal rating data, the above mentioned aspects should be considered sufficient to be acknowledged as the starting-point when evaluating the appropriateness of methods proposed in research publications as well as in books. To my knowledge, there have only been six research publications that to some extent discuss the properties of internal versus external rating data in connection to modelling rating migration matrices.[v] This is a minority regarding that there are forty or so research publications concerning credit ratings migration, all published during the 21st century.

First and foremost, there are two properties of the data to consider when founding an opinion on how to best model ratings migration. One of them is the first-order Markov property, i.e. that transition probabilities from each rating class are independent of the process history, also known as non-existing rating momentum or path dependence. That no rating momentum exists is very convening when modelling the transitions, since it simplifies the model. Although, it is a very strong assumption which adequacy could – and should – be argued, which is also the case in a number of research publications.[vi] In none of these publications, existence of path dependence in rating data can be rejected, even though it cannot always be proven to exist either.[vii] The majority however, find statistical differences that indicate that rating momentum exists, which also follows from logical arguments concerning the data characteristic. However, in one publication it was

---

[iii] Mählmann (2006) considers that (A) exact times of movement between rating classes and the class occupancy in between observations are unknown and (B) observation and inter-examination times vary for different borrowers. He argues that the observation time of the rating depends on (a) the current rating (reflects the probability of default, PD), (b) the terms of the loan (reflect the loss given default, LGD), (c) the loan size (reflects the exposure at default, EAD) and (d) the closeness of the bank–borrower relationship (reflects the moral hazard component of credit risk). Thus the structure of the monitoring times will vary for each borrower, depending on his level of expected loss (EL = PD x LGD x EAD).

[iv] Mählmann (2006) points this out in Basel Committee on Banking Supervision, 2005, Section 425.

[v] Araten et al. (2004), Feng et al. (2008), Lando & Skødeberg (2002) [consider both internal and external continuous, although without reflecting on what has been discussed in this text], Mählmann (2006), Rösch (2005) and Weissbach& Dette (2007)

[vi] Mählmann (2006), Parnes (2007), Frydman & Schuermann (2008), Bangia et al. (2002), Lando & Skødeberg (2002), Gagliardini & Gourieroux (2005)

[vii] Bangia et al. (2002)

found that the rating momentum effect of the previous move vanishes after about 2.5 years, but is significant up to that point in time.[viii]

The other property to be considered is the aspect of time-homogeneity, i.e. that the transition probabilities are functions of the distances in time between dates, and that the same transition matrix could be used for each point in time, independently of the date's position in time. This too is a strong assumption to impose, and rating data – at least external rating data – are shown to not have this property in numerous research publications,[ix] a property that is discussed in several others.[x] The fact that time-in-homogeneity most often has been shown for external data should be considered important, since external data often are ascribed to be through-the-cycle rather than point-in-time.[xi] Therefore, by logical argument, time-in-homogeneity – if existing – should be more pronounced when dealing with internal ratings, which have been shown to be more point-in-time than external ratings.[xii] However, for the sake of modelling ratings migration for a limited time frame, the important feature is not whether or not ratings possess time-homogeneity in an unlimited time frame, but if it approximately behaves like it for the time frame one is interested in.[xiii] The fact that this could be the case has been shown in research publications for time frames of 1-2 years.[xiv] However, this test was carried out for external rating data, which could reflect the stability of the through-the-cycle ratings. That this is the case is not contradicted by a publication that finds external rating data to exhibit a homogeneous process, and internal rating data to exhibit in-homogeneities.[xv] The aspect of limited time frames is likewise applicable on the first-order Markov property, although it is debateable whether the through-the-cycle rating has the same impact as it has on the homogeneity property. By logical argument, the difference in definition should – unlike in the case of time-homogeneity – be of no utmost importance.

Disregarding whether or not a model with first order Markov properties is possibly suitable, it has been shown that the use of discrete-time data has larger impact than imposing time-homogeneity on a continuous-time model. By comparison, it was found to be more damaging in terms of efficiency loss to use a so called discrete-time cohort method, than to use a continuous-time duration model with an assumption of time-homogeneity.[xvi] This, however, can only be considered shown for these specific models, and not representing a universal disparity between discrete-time and continuous-time methods. Overall, it is not foregone how to compare different models proposed by research publications due to their complex nature, and thereby to what extent they are affected by the properties of the data. This concerns both the properties the model claims are taken into account and – perhaps more importantly – those that are not.

[viii] Fledelius et al. (2004)

[ix] Kiefer & Larson (2007), Parnes (2007), Mählmann (2006), Hanson & Schuermann (2006)

[x] Jafry & Schuermann (2004), Bangia et al. (2002), Czado & Pflüger (2008)

[xi] Through-the-cycle ratings are forward looking, and as such nearly constant over time and not conditioned on the point of the economic cycle, whereas a point-in-time rating incorporates all relevant information that influences the one year worthiness of a borrower, Rösch (2005).

[xii] Mählmann (2006)

[xiii] When representing ratings migration as a Markov model, default is considered an absorbing state, which strictly speaking implies that – in the long run – all assets are in default. This fundamental property is not adequate to ascribe to assets in the long run, but appropriate in the short run, why Kiefer & Larson (2007) consider it "ridiculous" to worry about if the ratings really have the non-homogeneity property for an unlimited time frame.

[xiv] Kiefer & Larson (2007)

[xv] Weissbach & Dette (2007) applied their test on external rating data on sovereigns, which should be taken into consideration when assessing their results significance. Regarding the internal rating data they found in-homogeneities for few transitions to neighbouring rating classes.

[xvi] Jafry & Schuermann (2004) made a comparison between the discrete cohort method and one continuous-time duration model with an assumption of time-homogeneity and one without; between the latter two there were no statistical difference, but an efficiency gain was found in comparison to the cohort method.

Two final aspects regarding the data are the facts that the sample both is right-censored and left truncated, which concern internal and external credit ratings likewise. This is due to the fact that one does not know what happens to the firm after the sample window closes – e.g. does it default right away or does it live on until present – and that firms only enter sample if they have either survived long enough or received a rating.[xvii] These too are aspects that not too often are regarded in research publications,[xviii] mostly due to the fact that they are hard to do anything about.

When contemplating the characteristics of credit migration data and one tries to settle which characteristics that cannot be disregarded at any time, and which that under certain circumstances can be disregarded, there are things to keep in mind; that there are other aspects than the – perhaps simultaneous – dependencies of the transitions on time and credit ratings that has to be considered. If time-in-homogeneity is a property to be regarded when modelling ratings migration, one has to ponder how to best reflect this behaviour in the model. There are several research articles that evaluate the usage of macro economic variables in order to take the time-in-homogeneity into consideration.[xix] There are at least two clear discernable ways to do this; the model either condition the transition matrix on a macro economic variable considered relevant, e.g. the gross domestic product (GDP), interest rates or an economic cycle index etc,[xx] or there is a latent variable, which can be constructed in several different ways.[xxi] One of the latter actually back out what macro variable to be the most appropriate after the computations are done, and find that the GDP is the most suitable.[xxii] Over all, of all the made comparisons, the GDP is most often found significant, but regarding the economic cycle indexes there is no uniform conclusion.[xxiii] Other parameters are not compared to such extent that any conclusions can be drawn.

Another important aspect regarding macro economic variables, not commonly discussed in research papers, are the aspect of time lags. Only a couple of the articles applying macro economic variables reflect on this issue,[xxiv] why its importance has not been examined to an extent sufficient to draw any conclusions. Although, this has to be taken into consideration, since its impact by logical reasoning could be significant.

The aspect of including macro economic variables in the modelling of credit ratings migration is also reflected in the choice of discrete-time versus continuous-time modelling, an aspect not ever mentioned in literature. Macro economic variables are not observed on a continuous basis, why the modelling of such, if carried out, will contain a great uncertainness that will affect the modelling of transitions in general. Another aspect is the so to speak 'reason' that macro economic variables are not continuously observed; they often do not change that quickly. This further fortifies that continuous-time modelling regarding macro economic variables may be unsuitable, which in turn influences ones taken position on the discrete versus continuous aspect of the rating transitions themselves.

[xvii] Jafry& Schuermann (2004)
[xviii] Mählmann (2006), Jafry & Schuermann (2004)
[xix] Bangia et al. (2002), Parnes (2007), Koopman et al. (2008), Banachewicz et al. (2006), Duffie et al. (2007), McNeil et al. (2007), Trück (2008), Hanson & Schuermann (2004), Gagliardini & Gourieroux (2005), Feng et al. (2008), Nickell et al. (2000), Frydman & Schuermann (2008)
[xx] Bangia et al. (2002), Trück (2008), Frydman & Schuermann (2008)
[xxi] Koopman et al. (2008), Feng et al. (2008)
[xxii] Feng et al. (2008)
[xxiii] Note that these evaluations often are based on external, as well as American, rating data.
[xxiv] Czado & Pflüger (2008), Banachewicz et al. (2006), McNeil et al. (2007)

The conclusions one comes to regarding the above mentioned aspects are that time-in-homogeneity should be taken into account when modelling internal credit ratings migration, due to their point-in-time property. Rating momentum, on the other hand, could initially be disregarded. Regarding the truncation and censoring aspects, truncation can be disregarded, whereas censoring to some extent could – and should – be taken into consideration, if there is additional knowledge about whether a borrower has defaulted or not after the sample window closes. The fact that defaults are rather rare, and also that the default probability is the most interesting probability in the transition matrix, makes one wants to capture the most of the information available. The main points of decision, concerning the discrete-time versus continuous-time modelling, become the aspects of how to most adequately take macro economic variables into consideration, and how one assesses the monitoring of internal rating data. Due to the fact that these are to be taken into consideration simultaneously, there are other aspects than those most frequently discussed in research publications that are of decisive importance. This unexplored area will therefore have to be entered without extensive previous delving.

PROPOSITION OF A REFINED MATHEMATICAL METHOD FOR ESTIMATING TRANSITION PROBABILITIES

Taking the properties of internal rating data as well as macro economic data into consideration, it is deemed most appropriate in this thesis to model ratings migration with time-in-homogeneity properties as discrete in time, and to use a parametric approach to incorporate macro economic variables to reflect the state of the economy.[xxv] The model is then based on as few assumptions regarding the data as possible, which minimizes the risk that one incorporates incorrect assumptions with devastating result. As well, one makes use of all data available – which is important in view of the limited sample. Taking these aspects into account, a model is proposed with extraction in a continuous-time Maximum Likelihood estimation of a time-homogenous Markov chain, proposed by Mählmann (2006), which is suitable when dealing with incompletely observed rating data. The method is capable of providing ML-estimates when the exact time of movement between rating classes, and the class occupancy in between the observation times, are unknown – which is the gist other methods do not take into consideration. The method proposed in this thesis – the discrete-time ML-estimation – differs from the method proposed by Mählmann by being discrete in time at quarterly intervals, which is more suitable when incorporating macro economic variables. However, it still takes into consideration that the inter-examination times may vary for different counterparties, as well as for the individual. Furthermore, the discrete-time ML-estimation is then extended to incorporate macro economic variables, a feature not inquired by Mählmann.

TRANSITION PROBABILITY ESTIMATION WITH AN ASSUMPTION OF TIME-HOMOGENEITY

To deduce the ML-estimation methods and to show how the discrete- and continuous-time cases

[xxv] In line with previous reasoning, there will here on be an assumption of first-order Markov property. If one, however, do not want to disregard the aspect of rating momentum, it is easy to incorporate by expanding the transition matrix to include not only the initial rating and the end rating, but expanding the initial rating to also include the previous rating. By doing this one multiplies the number of elements in the matrix by the number of states, and therefore the number of elements to be estimated will be multiplied. All this, while the amount of data to base the computations on is the same. This will convey greater uncertainty in the ratings migration estimation. Due to the fact that estimation accuracy carries great weight in this thesis, in order to being able to compare methods, this rather simple evolvement to consider rating momentum will not be carried out, but can be considered for future research.

differ, a first-order time-homogenous Markov chain $X(t)$ denoting the rating class occupied at time t $(t \geq 0)$ serves as the starting point. Furthermore, the Markov chain assumes to have a finite number of states $1, 2, \dots, k_D$, where $1, 2, \dots, k_{D-1}$ are defined by decreasing levels of credit quality and k_D is the default state, which is absorbing. Let $P(s)$ denote the $k_D \times k_D$ transition probability matrix whose (kk')th element $p_{kk'}(s)$ is the probability of migrating from state k to k' within a time interval of length s. Assuming that the data are discrete and that the observation times of the Markov chain are identical for each borrower, i.e. $t_0 < t_1 < t_j < \cdots < t_{j_n}$ and equally spaced such that $s_j = t_j - t_{j-1} = s$ for all $j = 1, \dots, j_n$ monitoring steps, the ML-estimators of the stationary transition probabilities given by

$$\hat{p}_{kk'}(s) = \frac{N_{kk'}}{\sum_{k'=1}^{k_D} N_{kk'}}, \quad k, k' = 1, \dots, k_D \qquad \text{[I]}$$

are the *cohort* estimators, where $N_{kk'} = \sum_{j=1}^{j_n} N_{kk'}$ is the total number of recorded transitions from k to k'. The cohort estimators are the most commonly used in spite of well-known weaknesses; if observation times are not identical and equally spaced for all counterparties, the inadequacy of the cohort estimator is clear. This because the cohort estimator then is not the ML-estimation of $p_{kk'}(s)$, due to the fact that not all observations are used in this estimation – only those with identical and equal monitoring steps. In simple terms, the cohort approach just takes the observed proportions from the beginning of e.g. a year to the end of it as estimates of migration probabilities, which means that any observations within the year are not taken into consideration. The discrete-time ML method will incorporate these non-used observations. In order to elaborate on this, the continuous-time case is first considered.

The Markov chain in the continuous-time case introduces the concept of transition intensities $q_{kk'}(t, s) = q_{kk'}$. This means that given that class k is entered at calendar time t, and is still occupied at $t + s$, the transition out of k is determined by the set of $k_D - 1$ transition intensities $q_{kk'}$, with $q_{kk'} \geq 0$, $q_{kk} = -\sum_{k' \neq k} q_{kk'}$, let Q denote the $k_D \times k_D$ intensity matrix. For a time-homogenous Markov chain the relationship between the intensity matrix Q and the transition matrix $P(s)$ is given by

$$P(s) = e^{Qs} \qquad \text{[II]}$$

The ML-estimator of the transition intensity is in this case given by

$$Q_{kk'} = \frac{\sum_{i=1}^{n} N_{kk'}^{i}}{\sum_{i=1}^{n} T_k^i}, \; k' \neq k \qquad \text{[III]}$$

where $n_{kk'}^i$ denotes the number of $k \to k'$ transitions made by borrower i and T_k^i is the total time spent by i in class k.[xxvi] This so called *duration* estimator counts all rating changes over the course of e.g. a year, and divides by the total time spent in each rating – i.e. the exact transition times have to be known. If the data are discrete, in the sense that observations consist of the classes occupied by the borrowers at a sequence of discrete time points, the duration intensities cannot be used to estimate transition intensities. With no information about the timing of events between observations times, or about the exact transition time, neither the denominator nor the numerator can be calculated. Assuming that the data is discrete, one also must assume that the observation times are identical and equally spaced for the cohort estimator [I] to give the most probable estimations, i.e. the maximum likelihood.

[xxvi] For details, see Lando & Skødeberg (2002).

Mählmann (2006) outlines a method that provides ML-estimates of the intensity matrix Q. The method takes into consideration that exact times of movement between rating classes, and the class occupancy in between the observation times, are unknown and that observation times are assumed to be arbitrary. In order to derive the ML-estimates of the intensities, the key point is to formulate the transition probabilities $p_{kk'}$ in terms of intensities, i.e. it is the vector θ of intensities $q_{kk'}(k \neq k')$ that is to be estimated.[xxvii] Using a canonical decomposition and assuming $Q(\theta)$ has distinct eigenvalues $\lambda_1, \dots, \lambda_{k_D}$, from [II] one has

$$P(s) = A_0 \text{diag}\left(e^{\lambda_1 s}, \dots, e^{\lambda_{k_D} s}\right) A_0^{-1} \quad \text{[IV]}$$

where A_0 is a $k_D \times k_D$ matrix whose kth column is the right eigenvector for λ_k.[xxviii] Of the in general complicated function of $P(s)$ of θ it is possible to numerically obtain the eigenvalues making up A_0 and then $P(s)$. The individual contribution to the total likelihood function that is to be maximized is then divided in observations of counterparties that defaults during the observed time period $t_{i0} = 0, t_{i1}, \dots, t_{ij_i}$, with the rating of counterparty i as $x_{i0}, x_{i1}, \dots, x_{ij_i}$ respectively and $x \in \{k, \dots, k_D\}$, and those that are censored at the end of the observation time. Assuming that the exact time of default is known, but that the rating class on the previous instant before default is unknown, the censored observations[xxix] – known only to be a state in the set $R = \{1, \dots, k_D - 1\}$ with no additional information – contributes to the likelihood by

$$L_i(\theta) = \prod_{j=0}^{j_i-1} \left[p_{x_{ij} x_{ij+1}}\left(t_{ij+1} - t_{ij}\right) p^{\#}_{x_{ij_i}}\left(t_{ij+1} - t_{ij}\right) \right] \quad \text{[V]}$$

where for $k = 1, \dots, k_D - 1$,

$$p^{\#}_k\left(t_{ij+1} - t_{ij}\right) = \sum_{r \in R} p_{kr}\left(t_{ij_i+1} - t_{ij_i}\right) \quad \text{[V']}$$

The non-censored observations, i.e. those counterparties i who enter default during the observed period at t_{ij}, contributes to the likelihood function by

$$L_i(\theta) = \prod_{j=0}^{j_i-2} \left[p_{x_{ij} x_{ij+1}}\left(t_{ij+1} - t_{ij}\right) p^{¤}_{x_{ij_i} k_D}\left(t_{ij+1} - t_{ij}\right) \right] \quad \text{[VI]}$$

where for $k = 1, \dots, k_D - 1$,

$$p^{¤}_{k k_D}(t) = \sum_{k'=1}^{k_D-1} p_{kk'}(t-1) q_{k' k_D} \quad \text{[VI']}$$

The total likelihood function is then simply the product of the likelihood contributions over all n counterparties, conditional on the distribution of counterparties among the states at t_{i0}[xxx]

$$L(\theta) = \prod_{i=1}^{n} L_i(\theta) \quad \text{[VII]}$$

In order to transmit [VII] into the discrete-time case in quarterly intervals, notice that [II] can easily be rewritten as

---

xxvii Mählmann (2006)

xxviii Mählmann here leans on the work of Cox & Miller (1965).

xxix The censoring aspect can to some extent be taken into consideration if there is known fact of the survival status of the counterparties even after the original observation period is closed, which then quite easily can be included in the non-censored observations.

xxx For additional notes, see Mählmann (2006) pp 3246-3247.

$$P(s) = e^{Qs} = e^{Q\frac{n}{4}} = \left(e^{Q\frac{1}{4}}\right)^n = P_{quarter}{}^n \quad \text{[VIII]}$$

where $n = 1, \ldots., m$ quarters, e.g. the one-year transition matrix corresponds to $n = 4$, and $P_{quarter}$ is the one-quarter transition matrix.

In accordance with the line of reasoning for the continuous-time case, the contribution to the total likelihood function for counterparty i is

$$L_i(P_{quarter}) = \prod_{j=0}^{j_i-1} p_{kk'}{}^{(t_{ij+1}-t_{ij})} \quad \text{[IX]}$$

where $p_{kk'} = P_{quarter}(k, k')$, regardless if the counterparty i is in default at the end of the period or not. The total likelihood function can then be written as

$$L(P_{quarter}) = \prod_{i=1}^{n} \prod_{j=0}^{j_i-1} p_{kk'}{}^{t_{ij+1}-t_{ij}} \quad \text{[X]}$$

and the logarithmic likelihood function

$$\ln L(P_{quarter}) = \sum_{i=1}^{n} \sum_{j=0}^{j_i-1} (t_{ij+1} - t_{ij}) \ln p_{kk'} \quad \text{[XI]}$$

This, like in the continuous-time case, means that the one-quarter transition matrix is estimated with all available observations taken into consideration, and that only the exact time of default is assumed to be known.[xxxi] Thus, the rating class on the previous instant before default is unknown – i.e. one has to know when the counterparties have been evaluated, but not the exact time of transition from one rating class to another. Note that in the discrete-time case there is no need to iterative compute eigenvalues to obtain a solution, which makes the discrete-time ML method more straightforward. In order to obtain a solution to the ML-function in the general case of $k_D \geq 2$, a Quasi-Newton algorithm can be employed that uses finite differences to obtain numerical approximations of the derivatives.[xxxii] As the Quasi-Newton algorithm is an iterative process, a starting value is required. For this, the cohort-estimation will be used. The cohort-estimation will also be useful for comparative purposes, in order to establish a difference between the cohort estimation and discrete-time ML-estimation.

INCORPORATION OF THE STATE OF THE ECONOMY IN THE ESTIMATION OF TRANSITION PROBABILITIES

In order to extend the time-homogenous Markov chain estimated by the discrete-time ML method to incorporate the state of the economy, there is no extensive research to lean upon. By reflection, one can deduce two main lines of reasoning, which for simplicity are illustrated assuming only one macro economic variable.[xxxiii]

[xxxi] In this specification an 'exact' time is defined as the true quarter of the year that the default took place. This is a great strength of the discrete-time ML-estimation, in relation to the continuous-time case, in light of the properties of data discussed previously; even though a default can be assumed not to go unnoticed for long, it is still not foregone to know the exact date. In case of a default it is most probable that the counterparty is trying to stall the default, while hoping for a last minute salvation.
[xxxii] This is also the case with the continuous-time ML-estimation, Mählmann (2006).
[xxxiii] The line of reasoning can easily be extended to include any number of variables one may deem desirable.

[i] The ratings migration matrix with a time dependent factor is estimated in one step by parametrizing the macro economic variable by adding it to every row, with coefficient weights that add up to zero – i.e. each row still sums up to one after the macro economic variable is added.

[ii] Two time-homogenous ratings migration matrices are estimated conditioned on the state of the economy, i.e. one transition matrix is estimated on historically known times of expansion and one transition matrix is estimated on historically known times of recession. In the second step, the overall transition matrix estimated on all available observations is used as the average probabilities of transition. Depending on the macro economic variable, the average matrix will be shifted towards better or worse times by a relationship between the macro economic variable and the difference between the two state of the economy dependent matrices.

Before entering more deeply into the specifications of these methods to establish a transition matrix that incorporates the state of the economy, there are aspects even now distinguishable that are decisive when determining the most adequate way to proceed. In [i], one has to have a preconceived notion about how to distribute the weight among the elements in each row, i.e. one cannot let the data alone decide the specification of the function in an e.g. ML-estimation of the coefficients. While [i] therefore has the attraction of estimating the whole function at once, the drawback of having to specify the weights for each row outweigh the attraction. In [ii], one does not have to have a preconceived notion about the properties of the specification, other than what degree of relationship one wants to set – i.e. the data will on its own determine the most appropriate coefficients. Due to this, a specification in line with [ii] will be further developed and set up.

The first aspect to be considered is how to estimate matrices conditioned on the state of the economy. Conditioned matrices is a field that has been investigated in literature,[xxxiv] and the approach deemed most appropriate for this thesis purposes is described in Bangia et al. (2002). Bangia et al. condition their matrices on the business cycle and whether it can be considered to be in a recession or an expansion.[xxxv] These conditioned matrices are found to be significantly different. In order to obtain two matrices – conditioned on the state of the economy – one apply a chosen estimation method on two subsets of observations that reflect the desired states of the economy.

In order to specify a function in accordance with [ii], what remains is to establish the relationship between the macro economic variable and the economy-reflecting matrices. The most straightforward way is to assume a linear relationship between these, [xxxvi] i.e.

$$P_{quarter}(c) = P_{average} + (\alpha + \beta \times c)\left(P_{expansion} - P_{recession}\right) \qquad [XII]$$

where c is a macro economic variable and α and β are to be estimated. Note that $P(c)$ is a one-quarter transition matrix, and therefore the matrices part of the function are as well. The estimation of α and β is done by e.g. a ML-estimation, where the log-likelihood function, in accordance with the time-homogenous case [X], is

---

[xxxiv] Bangia et al. (2002), Trück (2008), Frydman & Schuermann (2008)

[xxxv] Bangia et al. (2002) use the cohort-estimation method and have only obtained external data, however one can conclude from previously discussed properties that their results are transferable to our circumstances.

[xxxvi] There is no limitation that requires a linear relationship, but since the aim of this thesis is to descry an adequate method to incorporate the state of the economy as a factor influencing the transition probabilities – and the line of reasoning deemed most appropriate has resulted in an inquire into a field previously not delved – it is found the most agreeable by way of introduction to make an as simple specification as possible.

$$L[P_{quarter}(c)] = \prod_{i=1}^{n} \prod_{j=0}^{j_i-1} p_{kk'}(c_j)^{t_{ij+1}-t_{ij}} \qquad [XIII]$$

with the same notation as previously and $p_{kk'}(c_j)^{t_{ij+1}-t_{ij}}$ the kk'th element of $P_{quarter}(c_j)^{t_{ij+1}-t_{ij}} = [P_{average} + (\alpha + \beta \times c_j)(P_{expansion} - P_{recession})]^{t_{ij+1}-t_{ij}}$, where c_j is the macroeconomic variable at the time t_{ij}. The reason for this specification of c_j is that at time t_{ij}, one cannot know the value of c at the future observations, why in this thesis the c at the 'current' time will be used in order to predict the future movements.

In the case where i macro economic variables are incorporated, [XII] is easily accommodated to

$$P_{quarter}(c_1, \dots, c_i) = P_{average} + \sum_{j=1}^{i}(\alpha_j + \beta_j \times c_j)(P_{expansion} - P_{recession}) \qquad [XIV]$$

Hereafter, in order to establish the appropriateness of the method, the focus will be the simple case with only one macro economic variable incorporated.

COMPARATIVE MEASUREMENT AND STATISTICAL DIFFERENCE

In order to statistically assess differences between ratings migration matrices, a bootstrap technique is applied in combination with a metric designed to measure the scalar differences between matrices. There are several appropriate metrics to apply of which the metric defined as the average of the singular values of the mobility matrix has been chosen.[xxxvii] A statistical difference by the bootstrap technique is achievable only in the time-homogenous case, although the metric, without confidence intervals, can always be obtained.[xxxviii] The bootstrap method is preferable to analytical confidence intervals due to the fact that these are not clear how to obtain for the ML-estimation methods.[xxxix] An important advantage with the bootstrap, in comparison with e.g. a Monte-Carlo approach, is that one does not have to have a preconceived notion about what distribution the sample originates from, i.e. the data speak all for itself.[xl]

The comparative measure of the average of the singular values of the mobility matrix

The metric as the average of the singular values of the mobility matrix, called *the singular value of decomposition metric* M_{SVD}, indicates what in simple terms can be expressed as the average amount of migration contained in a matrix.[xli] M_{SVD} is distribution discriminatory, which means that the metric discriminates between matrices having the same diagonal probabilities, but different off-diagonal distributions. This distinction between matrices with the same amount of mobility is important in the context of credit risk, since a migration to the far right or left of the transition

xxxvii The metric was proposed by Jafry & Schuermann (2004), who has conducted a thorough exploration and exploitation of transition matrices structure's to obtain a scalar metric that takes the properties of migration matrices into consideration. Their proposed metric is therefore preferable to metrics such as the cell-by-cell distance metrics, and eigenvalue and eigenvector based metrics. This metric was also used by Mählmann (2006), why it is suitable for comparative reasons.

xxxviii "A bootstrap based on resampling presumes that the data are serially uncorrelated or independent as the resampling process naturally reshuffles the data" Hanson & Schuermann (2006) p 2286. This is an assumption not easily transferred to the case of time-dependence.

xxxix In line with the reasoning in Hanson & Schuermann (2006).

xl This is also the case of the Jackknife resampling, where one has to add the assumption of the distribution in order to generate confidence intervals.

xli Jafry & Schuermann (2004)

matrix has significantly different consequences than near migrations in an economic and financial sense.[xlii] Given that the ratings migration matrix P is of dimension $k_D \times k_D$, and the identity matrix of the same I, and $\lambda_j(G)$ is the jth eigenvalue of a $k_D \times k_D$ matrix G, the metric is given by

$$M_{SVD}(P) = \frac{\sum_{i=1}^{k_D} \sqrt{\lambda_i([I-P]'[I-P])}}{k_D} \qquad \text{[XV]}$$

Note that only the dynamic part of the matrix P is left when the identity matrix is subtracted, which reflects the 'magnitude' of P in terms of implied mobility. The distance metric between the migration matrices P_I and P_{II}, preferably estimated with different methods, is then

$$\Delta M_{SVD}(P_I, P_{II}) = M_{SVD}(P_I) - M_{SVD}(P_{II}) \qquad \text{[XVI]}$$

The Bootstrap Method and Statistical Measurements

A non-parametric bootstrap is conducted by resampling the counterparties' rating histories, i.e. a sample is created by sampling with replacement counterparties with histories from the original sample until the bootstrap sample is of the same size as the original sample. However, there are two different ways to compare the size of a sample; the number of counterparties or the amount of rating history, i.e. the number of observation pairs. The latter is instinctual the most adequate way, since one is not interested in the counterparties as such, but their total rating history. Alas, there are great disadvantages when a sample is to be divided into subsamples, representing different times of economy, since there is no sensible way to affect the number of observation pairs in each subsample in this case.[xliii] Another aspect is that, due to the in general large number of counterparties, the irregular length of the rating histories does not result in large variations in the number of observation pairs anyway.[xliv] The slight variation of the number of observation pairs can also be considered to be a property with an intrinsic value, due to the fact that the data already have been established to not contain the complete rating transition history. Taking these aspects into consideration, it is deemed most appropriate to consider the sample size as the number of counterparties.[xlv]

In order to obtain the standard deviation for the estimated transition probabilities, and to obtain a confidence interval for the comparative metric M_{SVD}, a number of $B = 1000$ replications are conducted.[xlvi] In short, counterparties and their rating histories are picked randomly and replaced until the number of counterparties is the same as in the original sample. Then, cohort- and discrete-time ML-estimations are conducted. Then the resampling and estimations are repeated $B - 1$ times. For each estimation, the comparative metric is computed and the confidence interval of magnitude α, e.g. $\alpha =$ 95%, is obtained by sorting the comparative metrics in descending order and examining the breakpoints of the top and bottom $(1-\alpha)/2$ -percentiles, i.e. the lower boundary is the $[(1-\alpha)B/2\,]$th element of the vector and the upper boundary is

---

[xlii] Mählmann (2006)

[xliii] This is the case when estimating matrices representing economic expansions and recessions, since the subsamples have to originate from the same total sample.

[xliv] Hanson & Schuermann (2005) found the coefficient of variation – i.e. the standard deviation divided by the expected value – of the number of their equivalence to observation pairs to be just under 1%.

[xlv] This in contrast to Mählmann (2006), who kept the same number of observation pairs for each and every bootstrap sample. However, Mählmann (2006) had no need to consider the aspects of subsample properties.

[xlvi] In literature, 200 replications for standard errors and 1000 replications for confidence intervals are suggested, but also two or three fold the numbers are discussed. [Efron & Tibshirani and Andrews & Buchinsky]. The number of 1000 replications was in this case considered sufficient.

the $[(B-(1-\alpha)B)/2\,]$th element of the vector. The standard error σ of the transition probabilities is obtained by taking the standard deviation

$$\sigma_{jj} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{B}\left[x_i - \frac{\sum_{i=1}^{B} x_i}{B}\right]^2} \qquad \text{[XVII]}$$

of the ratings migrations elements $x_{jj} = x_i$ in turn.

EMPIRICAL APPLICATION

SAMPLE DATA

The sample data used to empirically evaluate the proposed method of discrete-time ML-estimation is here described. In order to apply the time-dependent function, a subsample corresponding to Swedish industry counterparties is extracted and a suitable macro economic variable that reflects the state of the economy is chosen.

Total Sample

The data come from the internal rating system of Skandinaviska Enskilda Banken AB, publ. (SEB) and consist of quarterly rating history over a ten year period. In addition, there is also information about the counterparties' sector and their country of origin.[xlvii] Counterparties of all sizes are chosen to be considered, i.e. not to draw a limit of e.g. a minimum annual turnover, in order to obtain the maximum amount of data. For comparative purposes, a five-position rating system is created by consolidating the internal risk grades.[xlviii] The new system also contributes to a higher degree of certainty in the transition probability estimations, due to the fewer elements to be estimated in the matrix, while the amount of data remains the same.[xlix] In the new rating system, grade 1 corresponds to the lowest and grade 5 to the highest degree of credit risk.

The data have been assembled quarterly, when the ratings of all presently existing counterparties have been either evaluated and set to an appropriate grade, or no evaluation has been conducted and the same grade as previously remains. A thorough evaluation is normally conducted once a year, with the starting quarter of when the counterparty was first rated, if there is no occurrence of a reason that implies otherwise.[l] The fact that not all counterparties have been under a profound evaluation every quarter is not a problem, however the fact that one cannot tell which

---

[xlvii] As a starting point, the dividing up of sectors done by Statistiska Centralbyrån (SCB) in Sweden according to SNI2007 and Kadam & Lenk (working paper) and Banachewicz et al. (2006) have been used. Further, counterparties have been chosen to be considered as belonging to the sectors Industrials, Service, Finance or Government. Further. A counterparty is also considered to be international or national, depending on if they have subsidiaries and/or mainly operate worldwide or not.

[xlviii] The number of grades treated in literature varies, but are usually below ten and above five and consolidated from a larger number of grades. Mählmann (2006), when applying the continuous-time ML-estimation method, used a six-position rating system, which while being the benchmark for this study is leading. Due to the properties of the internal data, it is not suitable to create a six-position rating system in the case of the SEB data, whereas a five-position rating system is created instead.

[xlix] This can also be interpreted as the amount of data to base the estimations on for each element is greater, since the data for each original rating class is added to the new class.

[l] A counterparty can have several 'first' ratings due to the fact that a counterparty can come and go in a sample; when it after a period of non-settlements – and no rating – enter an agreement, it receives a newly evaluated rating.

counterparties that have been thoroughly evaluated and which whose grades have been set solely on the notion of the previous grade, is a considerable problem. The reason is twofold: (i) the data now imply that the exact quarter of a transition is known, even when it is not and (ii) if there is no history of rating changes, one cannot be sure that the interpolated grades are always true.[li] The reason that these aspects are considered problems, originate from the fact that there is a lot of non-information masquerading as information in these interpolations, which significantly lower the value of the true – hidden – information. In (i), the knowledge of the latest evaluation – and not having to assume an exact transition time – is the strength of our proposed estimation method, but it is drowned by the noise of the interpolated grades. The same goes for (ii), even if it is highly probable that the rating grade – if evaluated – should have been the same. Still, it is assumptions of the data one should not have to make.

In order to manage the existence of concealed non-information, a decision to wash the data was taken in order to disclose the real information. There is no way of revealing all the true information and remove all the noise, but regarding the circumstances it is the best option available. The knowledge that a yearly evaluation, starting at the quarter of the first occurrence of a rating, is the basic principle means that a change in rating grade before the yearly evaluation is due to a special evaluation. When a special evaluation has been conducted, it is set as the new quarter for the yearly evaluation.[lii] This is the algorithm by which the original data sample is washed, and thereafter the new five-position rating system is applied.[liii] An effect of the washing is that counterparties rated with the same grade for only two, three or four quarters in a row contain no pair of observations, and are therefore of no use and excluded. The data sample after the washing consists of approximately 50 000 counterparties and contains approximately 200 000 observation pairs. Of these, 78% had an inter-examination time of one year, i.e. four quarters, 18.5% of less than one year and 3.5% of more than one year.[liv]

Subsample

When applying the time-dependent discrete-time ML-estimation method, a macro economic variable that relates to the counterparties as well as reflects the state of the economy is required. The sample consists of counterparties that cannot be considered a homogenous group, why a single macro economic variable cannot be assumed to relate to the whole sample. Therefore, a subsample consisting of only Swedish industry counterparties is engendered that, after washing, consists of approximately 4 000 counterparties.[lv] The washed sample contains approximately 25 000 observation pairs. Of these 79% had an inter-examination time of one year, 18.5% of less than one year and 2.5% of more than one year.

A suitable macro economic variable is considered to be *the capacity utilization in industry*, a percentage that measures how much of the available production capacity that is utilized in

---

[li] There is no guarantee that a reason for an earlier evaluation is actually discovered, and if a yearly evaluation is due soon it is possible that the evaluation is postponed until that time. It also occurs that a long-time counterparty, with a stable rating grade, is not in fact yearly evaluated, but even less frequently.

[lii] Corollary, special evaluations that do not result in a change in rating grade are unfortunately not captured.

[liii] In the case of a default, the counterparty is allowed to re-enter the sample after a period of not rated.

[liv] Only inter-examination times of the maximum length of five years – i.e. 20 quarters – have been allowed due to that the likelihood function results in high-degree terms that with the higher degree, the greater the risk of cancellation is, which will add to the uncertainty of the estimation. Therefore, a limit to inter-examination times have been set in accordance with what seems a reasonable time frame to be considering a company, without such changes made that it is not adequate to say that it is the same company in the beginning as in the end of the period – a company not rated for years is highly probable to have gone through major changes.

[lv] There is no easy way to distinguish between national and international counterparties in the data, but the majority is assumed to be considered national due to that national counterparties ought to be more common.

industry. The capacity utilization in industry (CUI) is not mentioned in literature regarding time-in-homogenous Markov models representing ratings migration, but in this thesis it is deemed superior to the commonly used GDP and other macro economic variables previously mentioned. This is because the CUI is done by queries, whereas e.g. the GDP is estimated and often revised years after. Another important aspect to consider is that the CUI can be assumed not to be lagged in relation to occurrence in time, but it might be lagged in relation to the occurrence of the actual time of a rating – e.g. a defaulted counterparty does not default immediately when its definite problems occur, but is lagged to some extent. This means that even if a counterparty is evaluated and set to a certain rating grade at the time, it is just the lagged rating that is observed, due to the inertia that is inherent to the information flow. Due to the aim of this thesis there will be no entering into this area, but it can be considered for future research to further investigate these aspects. [lvi]

In order to estimate two matrices, corresponding to times of economic expansion respective recession, two additional subsamples have to be engendered by parting the original subsample in accordance with what the macro economic variable indicates. There are several varieties of a straightforward way of constructing threshold values of the macro economic variable; to calculate the average, or the median, value of the macro economic variable during the observed period and interpret the 'better' part of the sample as times of expansion, and vice versa with the 'worse' part. Another is to take the limits of the top respective the lowest thirds of the macro economic values as the threshold values.[lvii] Due to the limited amount of data, it is preferable to use as much of the available data as possible, and therefore the threshold value is chosen to be set to the median.[lviii] Due to the fact that some of the counterparties do not end up with a pair of observations when the sample is parted in time, the expansion subsample consists of approximately 4 000 counterparties and contains approximately 13 000 observation pairs. The recession subsample consists of approximately 3 000 counterparties and contains approximately 9 000 observation pairs. Of the expansion subsample, 55.5% of the observation pairs had an inter-examination time of one year, 18% less than one year and 26.5% of more than one year. The corresponding numbers of the recession subsample was 48%, 15% and 37%. There is also a considerable increase in number of defaults during the times defined as recessions, which indicates that an adequate macro economic variable has been chosen.

ESTIMATION OF TRANSITION PROBABILITIES

An empirical application of the methods proposed is conducted on the sample data as presented above.[lix]

Time-homogeneity

By maximizing the logarithmic likelihood function [XI] using all available ~200 000 rating

---

[lvi] Since the aim of this thesis is to establish an improved method, it can be considered for future research to further investigate the impact and suitability of other macro economic variables, such as the GDP. Also the case when several macro economic variables are incorporated and the possible impact of lags should be investigated, but for this thesis the purpose is to investigate the method and therefore an as 'simple' macro economic variable as possible is suitable.

[lvii] It is also possible to not only use the macro economic values corresponding to the observation period when setting the limits, but all available data – SCB started to record the CUI in 1980 – but in this case the amount of rating data are not extensive enough to result in enough observation pairs in each subsample for this to be considered.

[lviii] The median is chosen over the average due to the limited amount of data and the desire to reduce the effect of possible outliers.

[lix] All estimations and related work were implemented in MATLAB®.

observation pairs, a one-quarter transition matrix was estimated. Taking the one-quarter transition matrix to the fourth power, a one-year transition probability matrix is obtained.[lx] The one-year cohort estimator and the discrete-time ML-estimator, with standard errors obtained by a non-parametric bootstrap, is reported in Panel I and Panel II of Table I. The delta of the singular value of the decomposition metric is calculated with statistical significance at the 1% level[lxi] to 0.0781 with a standard error of less than 0.001, which can be considered to represent a substantial difference between the two matrices.[lxii]

Table I
Transition probabilities with different estimation methods

From	To 1	2	3	4	5	Default
Panel I: The one-year cohort estimated transition matrix						
1	**0.97779**	0.018416	0.0025728	0.0012187	0	0
	0.0017376	*0.0015776*	*0.00061017*	*0.00039946*	*0*	*0*
2	0.012286	**0.93989**	0.035834	0.010385	0.0011701	0.00043879
	0.00095821	*0.0020357*	*0.0015721*	*0.00085825*	*0.00030085*	*0.00018342*
3	0.00028629	0.0049724	**0.95707**	0.033662	0.0034355	0.00057258
	6.8131e-05	*0.00027831*	*0.00080998*	*0.00070823*	*0.00023209*	*9.0771e-05*
4	0.00025675	0.00071619	0.023296	**0.96023**	0.012445	0.0030539
	5.7238e-05	*0.00010323*	*0.00055866*	*0.00070511*	*0.00040707*	*0.00020531*
5	0.0002553	8.5099e-05	0.0041699	0.053698	**0.92384**	0.017956
	0.00014735	*8.3929e-05*	*0.00058472*	*0.0021015*	*0.002483*	*0.0012283*
Default	0	0	0	0	0	**1**

[lx] The one-year cohort estimator was used as the starting value for the iterative process due to the fact that there were not enough data to obtain reasonable one-quarter estimations. Observation pairs of a period of four quarters were used for the one-year cohort estimation and overlapped observations, i.e. observation pairs that start and/or ends during another observation pair's four quarter period, were allowed due to the time-homogeneity assumption, and that one wants to use as much available data as possible. However, the likelihood function turned out not to be convex, why a good starting value is a necessity. Therefore also the fourth root of the one-year cohort estimation, as well as the average of the one-year and the one-quarter cohort estimated matrix, were considered as starting values. Empirically it turned out that the one-year cohort estimation always found the best optimum. Alas, even the best local optimum turned out to be flat, and therefore the absolute optimum solution cannot be established. Even if one therefore cannot be certain of finding either the global optimum or the definitely 'best' probabilities, one always finds a better estimation than the available cohort.

[lxi] The 99% confidence interval was calculated to 0.0758 – 0.0810 and obtained by a non-parametric bootstrap as outlined above.

[lxii] As a comparative value, Mählmann (2006) calculated a statistically significant at the 5% level delta of the singular value of the decomposition metric to 0.00047 when comparing a cohort and continuous-time ML-estimation.

Panel II: The one-year discrete-time ML-estimated transition matrix

1	**0.94114**	0.046932	0.0081713	0.0031352	0.00027286	0.00034762
	0.0025518	*0.0022415*	*0.00088766*	*0.0005314*	*0.00013313*	*0.0002483*
2	0.033982	**0.85942**	0.081308	0.022454	0.0023591	0.00047481
	0.0013912	*0.0027947*	*0.0021078*	*0.0010787*	*0.00032629*	*0.00015149*
3	0.0013128	0.01448	**0.87685**	0.099323	0.0065818	0.0014561
	0.00014644	*0.00047353*	*0.001283*	*0.0011107*	*0.00027686*	*0.00012497*
4	0.00062057	0.0023687	0.069051	**0.88323**	0.037119	0.0076108
	8.6174e-05	*0.0001725*	*0.00091402*	*0.0011862*	*0.00066437*	*0.000294*
5	0.00046421	0.00056183	0.010031	0.15295	**0.76534**	0.070657
	0.00017942	*0.00017905*	*0.00069527*	*0.003101*	*0.0039626*	*0.0021599*
Default	0	0	0	0	0	**1**

Table I *Panel I* presents the cohort estimated one-year transition matrix, based on all yearly observation pairs, and *Panel II* presents the discrete-time ML-estimated one-year transition matrix, based on all observation pairs with an inter-examination time of less than or equal to five years. Note that the ML-estimated one-year transition matrix was obtained by raising the estimated one-quarter transition matrix to four. The diagonal elements are bolded for convenience and the standard errors obtained by a non-parametric bootstrap are in italics.

With the naked eye, note the important difference that the discrete-time ML method generated probabilities for all transitions at a one-year interval,[lxiii] whereas the cohort method estimates two transition probabilities to zero in the one-year transition matrix. Also, only the transition probability from 1 to 5 in the ML-estimated transition matrix is marginally below the limit of 0.03% that is set by the Basel accords. Further, the cohort estimated transition matrix is considerably more diagonal dominant than the ML-estimated transition matrix, which indicates that the ML-estimation is better capturing mobility. Most importantly, the methods generate different default probabilities. For all rating grades, the less efficient cohort method underestimates the default risk. This is due to the fact that the one-year cohort estimator only uses the yearly noted transitions, whereas the ML-estimation makes a difference between events that occurred during the year, and those that actually took place at the end of the year.[lxiv] Note that the ML-estimation captures the greater uncertainty of remaining in the highest degree of risk for more than a year; the probability of remaining in the same rating class, i.e. 5, is more than ten percentage points less than every other probability of keeping the same grade. In fact, the possibility of recovering from the worst rating grade, i.e. transfer from 5 to 4, is much higher – 15.3% – than for any other grade, even in the cohort estimated transition matrix. When considering the created five-positions rating system, this is probably due to that the risk classes

[lxiii] Note that non-zero estimates also were obtained at the one-quarter estimation.

[lxiv] Due to that only yearly observation pairs were used to conduct the cohort estimation, defaults that did not occur on a one-year basis since a recorded rating grade are not taken into account due to that they are then set to not-rated. In the original sample, there was no standardized time frame before the defaulted counterparty was set to not-rated, why it sometimes would have been taken into account and sometimes not. One may argue that once a counterparty has entered default it should stay there and therefore be recorded by the cohort-estimator, which was the case in Mählmann's study. Due to that definition of default, he instead got an overestimation in default risk for the cohort estimator, see Mählmann (2006) p 3247. The cohort estimator therefore can be concluded to either over- or underestimate the default probabilities.

translated into grade 4 are monitored closely enough to register the more volatile nature of these grades.

Time-dependence

In order to write out a time-dependent function, the subsample of the Swedish industry counterparties is used and time-homogenous average and expansion respective recession matrices are estimated. The one-year average transition matrices estimated with the cohort respective the discrete-time ML method, with standard errors obtained by a non-parametric bootstrap, is reported in Panel I and Panel II of Table II. The delta of the singular value of the decomposition metrics is calculated to 0.0983 with a standard error of 0.005, and is statistically significant at the 1% level,[lxv] representing a considerable difference. This average transition matrix calculated on the subsample presents the same characteristics as the average transition matrix estimated on the total sample, i.e. less diagonal-domination – and therefore capturing greater mobility – as well as non-zero estimates of the one-year transition probabilities. The one-year transition matrices for times of expansion respective recession, estimated with the discrete-time ML method, are presented in Panel I and Panel II of Table III, with standard errors obtained by a non-parametric bootstrap. A statistically significant delta of the singular values of decomposition at the 6% level is calculated to 0.0165 with a standard error of 0.008,[lxvi] a substantial difference that represents the capturing of diverse states of the economy.[lxvii]

Table II
Transition probabilities with different estimation methods, subsample

From	To 1	2	3	4	5	Default
Panel I: The one-year cohort estimated transition matrix						
1	**0.92523**	0.037383	0.03271	0.0046729	0	0
	0.019004	*0.012913*	*0.013439*	*0.0046912*	*0*	*0*
2	0.0026397	**0.93797**	0.045315	0.012319	0.0017598	0
	0.0011007	*0.0051492*	*0.0044375*	*0.0023512*	*0.00087969*	*0*
3	0.00047744	0.0056099	**0.94891**	0.040224	0.0041776	0.0005968
	0.0002394	*0.00082724*	*0.002472*	*0.0021569*	*0.00070102*	*0.00027279*
4	0	0.00047642	0.031324	**0.95248**	0.014173	0.0015484
	0	*0.00023634*	*0.0019968*	*0.0023132*	*0.0013272*	*0.0004261*
5	0	0	0.004085	0.073529	**0.91422**	0.0081699
	0	*0*	*0.0018907*	*0.0075156*	*0.0082814*	*0.0025211*
Default	0	0	0	0	0	**1**

[lxv] The 99% confidence interval was calculated to 0.0854 – 0.1123 and obtained by a non-parametric bootstrap as outlined above.
[lxvi] The 94% confidence interval was calculated to 0.0007 – 0.0325 and obtained by a non-parametric bootstrap as outlined above.
[lxvii] As a comparative value, Jafry & Schuermann (2004) – who proposed the metric – calculated a delta of the singular value of the decomposition metric of 0.0434, when comparing transition matrices estimated with the parametric time-homogenous duration method for expansion and recession periods based on data from Standard & Poor from 1981-2002. The greater amount of data is probably the reason for the slightly greater difference when comparing the two duration estimated matrices.

Panel II: The one-year discrete-time ML-estimated transition matrix

1	**0.82947**	0.12928	0.01861	0.022099	0.00049731	4.3073e-05
	0.025422	*0.022574*	*0.007986*	*0.0084107*	*0.00015686*	*1.5988e-05*
2	0.0068321	**0.85372**	0.11219	0.024973	0.0021779	0.00011101
	0.0015362	*0.0068367*	*0.0059682*	*0.0026489*	*0.00078691*	*2.7768e-05*
3	0.0015719	0.018074	**0.86027**	0.1111	0.0080123	0.00096913
	0.00044153	*0.0014504*	*0.0038642*	*0.0033361*	*0.0008093*	*0.00029418*
4	0.00054665	0.0014436	0.080927	**0.87405**	0.038661	0.0043665
	0.00022138	*0.00032102*	*0.0030474*	*0.0036627*	*0.0020035*	*0.00064699*
5	5.2635e-05	0.00069625	0.0090504	0.19902	**0.73415**	0.057027
	2.0313e-05	*0.00063881*	*0.0012721*	*0.010494*	*0.012929*	*0.0059341*
Default	0	0	0	0	0	**1**

Table II *Panel I* presents the cohort estimated one-year transition matrix, based on all yearly observation pairs of the subsample, and *Panel II* presents the discrete-time ML-estimated one-year transition matrix, based on all observation pairs with an inter-examination time of less than or equal to five years in the subsample. Note that the ML-estimated one-year transition matrix was obtained by raising the estimated one-quarter transition matrix to four. The shaded elements are those where there are too few observations to obtain a one-quarter probability estimate above the lower limit of 10^{-6}. The diagonal elements are bolded for convenience and the standard errors obtained by a non-parametric bootstrap are in italics.

Table III
The discrete-time ML method applied to the subsamples representing expansion and recession

From	To 1	2	3	4	5	Default
Panel I: The one-year ML-estimated transition matrix of times of expansion						
1	**0.85567**	0.11091	0.0065303	0.026248	0.00062456	2.1212e-05
	0.027606	*0.025096*	*0.0049772*	*0.012238*	*0.00024168*	*9.2359e-06*
2	0.0058386	**0.84541**	0.11506	0.030236	0.0033856	6.3269e-05
	0.0018934	*0.009193*	*0.0086456*	*0.0043707*	*0.0013988*	*2.3483e-05*
3	0.00081461	0.017959	**0.85223**	0.12339	0.0054712	0.00013058
	0.00034726	*0.001851*	*0.0050249*	*0.0045594*	*0.00099826*	*4.268e-05*
4	0.00057791	0.0017902	0.083545	**0.8703**	0.041917	0.0018647
	0.00027389	*0.00044991*	*0.0037792*	*0.0045834*	*0.0026469*	*0.00046351*
5	6.2601e-05	0.0010733	0.0085225	0.22508	**0.73243**	0.032828
	2.8605e-05	*0.00094745*	*0.00069059*	*0.014547*	*0.016274*	*0.0057104*
Default	0	0	0	0	0	**1**

Panel II: The one-year ML-estimated transition matrix of times of recession

1	**0.78314**	0.1842	0.016536	0.015657	0.0004303	3.3314e-05
	0.036725	*0.03353*	*0.0088721*	*0.010328*	*0.00023157*	*1.9616e-05*
2	0.0072018	**0.84726**	0.12094	0.022539	0.0019597	0.00010395
	0.0021932	*0.0087858*	*0.0081867*	*0.0035732*	*0.001181*	*3.9097e-05*
3	0.0015069	0.017845	**0.86104**	0.10682	0.011707	0.0010775
	0.00053616	*0.001816*	*0.0051788*	*0.004508*	*0.0015068*	*0.00038422*
4	0.00030205	0.00068131	0.084445	**0.86646**	0.043444	0.0046662
	0.0002341	*0.00019159*	*0.0048548*	*0.0058968*	*0.0033622*	*0.0010095*
5	3.2818e-05	7.7602e-05	0.013446	0.20668	**0.72007**	0.059687
	2.2881e-05	*3.7025e-05*	*0.0040448*	*0.016292*	*0.019431*	*0.0090354*
Default	0	0	0	0	0	**1**

Table III *Panel I* presents the discrete-time ML-estimated one-year transition matrix, based on all observation pairs with an inter-examination time of less than or equal to five years in the expansion subsample. *Panel II* presents the discrete-time ML-estimated one-year transition matrix, based on all observation pairs with an inter-examination time of less than or equal to five years in the recession subsample. Note that the ML-estimated one-year transition matrix was obtained by raising the estimated one-quarter transition matrix to four. The shaded elements are those where there are too few observations to obtain a one-quarter probability estimate above the lower limit of 10^{-6}. The diagonal elements are bolded for convenience and the standard errors obtained by a non-parametric bootstrap are in italics.

Unfortunately, four probabilities of the average one-quarter transition matrix could not be estimated, and were therefore set to the lower limit of 10^{-6},[lxviii] see Table II. This was the case for seven respective six probabilities for the expansion and recession matrices, see Table III. Out of these, three probabilities of the average one-year transition matrix, and four respectively out of the expansion and recession matrices, were below the limit of transition probabilities set by the Basel accords to 0.03%. The reason that these probabilities could not be estimated is because the amount of data for these transitions was not enough to generate a probability, i.e. the transition did not occur. This belief was further supported by the fact that there were only four non-estimated probabilities for the estimation founded on the greatest amount of data – ~25 000 observation pairs – and three respectively two more for the estimations based on less, i.e. ~13 000 observation pairs for the expansion subsample and ~10 000 observation pairs for the recession subsamples.[lxix] In comparison, note that the one-quarter transition matrix estimated on all ~200 000 observation pairs did not result in a single probability set to the lower limit.

The statistically significant differences in default probabilities for the expansion and the recession ratings migration matrices – the default-probabilities are all higher for the recession transition matrix – indicate that our chosen macro economic variable CUI is an adequate variable to condition on, in order to obtain transition matrices reflecting different states of the economy. Also, note that the probability of remaining at the lowest risk grade 1 is significantly larger for the

[lxviii] Due to that there is no logarithm of zero, the probabilities are not allowed to smaller than a set limit, which in this case was set to 10^{-6}.

[lxix] The fact that the recession matrix had fewer elements set to the lower limit than the expansion matrix, in spite of the lesser amount of data, is due to the distribution of observation pairs and not the number of pairs as such, even though the greater number of observations pairs, the more likely observations of rarer transitions.

expansion transition matrix than for the recession transition matrix. Other non-diagonal transition probabilities indicate more ambivalence; only for grades 4 and 5 are there in total greater probability of an improvement than deterioration in credit risk, for the expansion transition matrix in comparison to the recession transition matrix. Overall, three out of five risk classes behave according with expectations, which is found satisfactory to hold to our specification conditioning on the macro economic variable CUI. Since there is no way of obtaining a prediction interval when applying the method that takes the state of the economy into consideration, these results are deemed sufficient when establishing the appropriateness of our specification.

By maximizing the likelihood function [XIII], using all available ~25 000 rating observation pairs, a value of $\alpha =$ -5.773 and $\beta =$ 0.0652 were estimated. The maximum and minimum historically recorded values of c were used as the upper and lower values c is allowed to take, i.e. c can be set to any value between 72.9% and 92%.[lxx] The minimum and maximum values have been set in accordance to all available data from 1980-2008, but if one deems other values should be included, the bounds can easily be expanded.[lxxi] The time-dependent one-quarter transition matrix is thus determined by

$$P_{quarter}(c) = P_{average} + (-5.773 + 0.0652 \times c)\left(P_{expansion} - P_{recession}\right) \qquad \text{[XII]}$$

with the previous notation. In order to appraise the time-dependent function, an application of empirical scenarios is conducted. In Table IV Panels I, II, III, IV, V and VI, six different historical scenarios are constructed of mediocre as well as more extreme times, reflecting the state of the economy, in order to estimate one-year transition matrices representing such times. The deltas of the singular value of the decomposition metric for these one-year matrices are presented in Table V. The delta values are of similar size in comparison to the previously established statistically significant delta values.

Table IV
The discrete-time ML method applied to the subsamples representing expansion and recession

From	To 1	2	3	4	5	Default
Panel I: The one-year ML-estimated transition matrix of average times						
	c	88.8	88.8	88.8	88.8	
1	**0.83049**	0.12825	0.018477	0.022248	0.00049939	4.2911e-05
2	0.0068131	**0.85368**	0.1121	0.025079	0.0021979	0.00011069
3	0.0015624	0.018075	**0.86014**	0.11133	0.0079248	0.00095578
4	0.00055055	0.0014589	0.080916	**0.87409**	0.038639	0.0043282
5	5.3021e-05	0.00071004	0.0089816	0.19928	**0.73433**	0.056656
Default	0	0	0	0	0	**1**

[lxx] A restriction to c has to be set in order to make sure $P(c)$ fulfils the requirements of a transition matrix, i.e. in this case that the elements of each row cannot be set to negative values. The fact that each row has to sum to one is self-fulfilled by the definition of the function.

[lxxi] When testing for different values of c, it turned out that if increasing the interval, α too increased in value, while β decreased.

Panel II: The one-year ML-estimated transition matrix of combined average times

	c	91.9	88.8	88.8	72.9	
1	**0.82265**	0.13692	0.018934	0.020953	0.00049913	5.3247e-05
2	0.006921	**0.85389**	0.11286	0.024052	0.0021187	0.00013966
3	0.0016261	0.018043	**0.86122**	0.1093	0.0087356	0.0010579
4	0.00052201	0.0013413	0.081135	**0.87346**	0.038839	0.0046874
5	4.5675e-05	0.00058874	0.0096078	0.19768	**0.73311**	0.058978
Default	0	0	0	0	0	**1**

Panel III: The one-year ML-estimated transition matrix of combined average-and-expansion times

	c	91.9	91.9	88.8	88.8	
1	**0.8379**	0.12093	0.017327	0.023284	0.00051883	4.4504e-05
2	0.0066624	**0.8535**	0.11155	0.025787	0.0023672	0.00011674
3	0.00149	0.01808	**0.85929**	0.11293	0.0073335	0.00085957
4	0.00057968	0.0015706	0.080867	**0.8744**	0.038492	0.0040719
5	5.4461e-05	0.00080568	0.0085104	0.20125	**0.73562**	0.053765
Default	0	0	0	0	0	**1**

Panel IV: The one-year ML-estimated transition matrix of combined average-and-recession times

	c	88.8	88.8	72.9	72.9	
1	**0.80776**	0.1507	0.022026	0.019043	0.00043509	3.7091e-05
2	0.0072755	**0.85424**	0.11381	0.022886	0.0016755	8.9615e-05
3	0.001785	0.018065	**0.86277**	0.10635	0.0097536	0.00126
4	0.00046115	0.0011123	0.081068	**0.87313**	0.039093	0.0051237
5	4.8869e-05	0.00041503	0.010439	0.19317	**0.73035**	0.065592
Default	0	0	0	0	0	**1**

Panel V: The one-year ML-estimated transition matrix of times of expansion

	c	91.9	91.9	91.9	91.9	
1	**0.8454**	0.11307	0.01654	0.024424	0.0005281	4.0154e-05
2	0.0065351	**0.85332**	0.11092	0.026631	0.00249	0.00010502
3	0.0014231	0.018098	**0.85835**	0.1147	0.0066498	0.00076406
4	0.00060801	0.0016822	0.080738	**0.87486**	0.038323	0.0037713
5	5.8831e-05	0.00091179	0.0079745	0.203	**0.73685**	0.051218
Default	0	0	0	0	0	**1**

Panel VI: The one-year ML-estimated transition matrix of times of recession

	c	72.9	72.9	72.9	72.9	
1	**0.78561**	0.17396	0.024459	0.015544	0.00038842	4.6585e-05
2	0.0076509	**0.85485**	0.1158	0.020258	0.0013036	0.00011547
3	0.0019842	0.01802	**0.8657**	0.10083	0.01186	0.001584
4	0.00037741	0.00076251	0.08147	**0.8717**	0.039612	0.0060617
5	3.7006e-05	9.1896e-05	0.012091	0.18781	**0.72658**	0.073406
Default	0	0	0	0	0	**1**

Table IV *Panel I-VI* present discrete-time ML-estimated one-year transition matrices, based on all observation pairs with an inter-examination time of less than or equal to five years in the subsample, with the value of the macro economic variable CUI c set to four different values, representing different scenarios during a year. In order of presentation there is an overall average, a combined average, a combined average-and-expansion, a combined average-and-recession, an overall expansion and an overall recession scenario. Note that the ML-estimated one-year transition matrix was obtained by taking the product of the four estimated one-quarter transition matrices representing a scenario. The diagonal elements are bolded for convenience, but no standard errors can be obtained in the case of time-in-homogeneity.

Table V
The deltas of the single value decomposition metric for the matrices in Table IV Panel I-VI

Delta *MSVD* Average – Expansion	-0.0033682
Delta *MSVD* Average – Recession	0.01052
Delta *MSVD* Expansion – Recession	0.013888
Delta *MSVD* Combined Average - Combined Expansion	-0.0035003
Delta *MSVD* Combined Average - Combined Recession	0.0034053
Delta *MSVD* Combined Expansion - Combined Recession	0.0069056
Delta *MSVD* Combined Average – Average	-0.0018
Delta *MSVD* Combined Average – Expansion	-0.0052
Delta *MSVD* Combined Average – Recession	0.0087
Delta *MSVD* Combined Expansion – Average	0.0017
Delta *MSVD* Combined Expansion – Expansion	-0.0017
Delta *MSVD* Combined Expansion – Recession	0.0122
Delta *MSVD* Combined Recession – Average	-0.0052
Delta *MSVD* Combined Recession – Expansion	-0.0086
Delta *MSVD* Combined Recession – Recession	0.0053

Table V The deltas of the singular values of decomposition metric for the different scenarios. For convenience, the notations of the combined average-and-recession and the combined average-and-expansion transition matrices have been abbreviated. Note that they are all in the same size as the previously established statistically significant deltas, and that the largest deltas are obtained when a recession scenario is a part of the difference.

Note the difference between the two average one-year transition matrices; the one constructed of only the average value has for all risk classes, except the lowest, a lower degree of default risk than the average transition matrix, constructed with a combination of average times and one period of expansion and one of recession. This indicates that stability is rewarded, which is in line with common sense regarding default risk. When comparing the risk of default from risk class 5 of all six matrices representing different scenarios,[lxxii] the probability of default increases for the different scenarios. The lowest risk is for the overall expansion transition matrix followed by the combined average-and-expansion transition matrix, the overall average transition matrix, the combined average transition matrix, the combined average-and-recession transition matrix and the overall recession transition matrix. This result is in line with sensibleness; the worse and/or unstable the state of the economy is, the greater is the risk of default.

When comparing the deltas of the singular value of the decomposition metric for these ratings migration matrices, the largest deltas are the ones for the overall average vs. overall recession transition matrices, the overall expansion vs. overall recession transition matrices and the combined average-and-expansion vs. overall recession transition matrices. These are the three transition matrices one, by common sense, expects to find the largest difference between. Note that the combined average-and-recession vs. the overall expansion matrices and the combined average vs. the overall recession matrices, are only slightly smaller. The slightest differences are between the combined average vs. average transition matrices, the combined average-and-expansion vs. expansion transition matrices and the combined average-and-expansion vs. the overall average transition matrices. The rest of the deltas are in absolute values of the same size, which is expected due to the construction of the scenarios,[lxxiii] which indicate that the differences in scenarios are mirrored in the estimations, and therefore the time-dependent function ought to be useful in stress tests.

The small difference between the average and combined-average transition matrices is expected, but not the small difference between the transition matrices in differently composed scenarios of expansion. This difference is also represented by the fact that even in the scenario representing the best state of the economy, the estimated ratings migration matrix has a higher default risk from risk class 5 than the time-homogenous estimated expansion transition matrix. This indicates that the time-dependent function is over-estimating the probabilities of default in times of expansion. The reason for this behaviour of the time-dependent function might be connected to that the recession transition matrix' default probabilities are only slightly larger than the average in the time-homogenous case. Since one above all is interested in transition probabilities in times of recession, it is preferable that the probabilities of default are over-estimated in times of expansion, to under-estimated in times of recession. Due to this, it is not considered a problem for using the current specification for stress testing.[lxxiv]

CONCLUDING REMARKS

As recounted, in the world of banking and finance credit ratings migration that measures the expected changes in credit quality of counterparties is of the greatest importance. The most commonly applied methods have proven inefficiencies and the trend in research is to propose even more complex methods, which due to their intricate nature are hard to grasp as well as to

---

[lxxii] The comparison of default transitions is restricted to the transition from 5 to default since these are most probably the best estimated default probabilities, due to the greater amount of data for defaults from risk class five than from any other risk class.

[lxxiii] The scenarios are representing states of the economy that are 'equally' distributed over historical times of expansion, recession and four states in between.

[lxxiv] If one for some reason wants to keep the transition probabilities between the values estimated in the time-homogenous case, the specification of the time-dependent function can be changed, e.g. replace the average matrix with the expansion matrix. The evaluation of other such specifications can be considered for future research.

implement. Moreover, the need of a method that takes the state of the economy into account – in order to render stress testing of not yet historically known scenarios – is great, while hitherto there is no such method recognized.

The method of ML-estimation in discrete-time proposed in this thesis rectifies many of these predicaments. As shown in this thesis, it is straightforward in theory as well as in implementation and attends to the shortages of the commonly used methods, as it uses all available sample data, while not imposing preconceived notions about their properties. It is an important feature due to the unfortunately often-limited sample, with characteristics seldom acknowledged. This is mirrored in that the discrete-time ML-estimated transition matrices differs significantly from the ones estimated with the cohort method, which only uses a fraction of the available information. The ML-estimated matrix is also much less diagonal dominant, which indicates that the movements even in less frequently observed transitions are better captured. The ML method also renders non-zero probability estimates for all one-year transition matrices.

The forthright and flexible way to implement macro economic variables as presented in this thesis has not been seen in literature before, and due to the directness of the method one is not deceived to impose a too great belief in the mathematics. But still, the method captures the extreme stressed movements one hoped to generate, and even between only slightly different scenarios a difference was discerned. The incorporation of macro economic variables opens for the possibilities of stress testing ratings migration matrices, with historically as well as fictive scenarios, one of the aims of this proof of concept.

The discrete-time ML-estimation method ought to supersede the cohort method in the case of time-homogeneity. In the field of further work in evolving the incorporation of the state of the economy, the use of several macro economic variables and their appropriateness for different subsamples should be investigated. Also the specification of the time-dependent ML-function can be further inquired into, as also the aspect of rating momentum, which also is of interest in the case of time-homogeneity.

ACKNOWLEDGEMENTS

Slutet på detta examensarbete är inte bara slutet på en tämligen omfattande mängd skrivna blad – vilka jag för övrigt hoppas att de som kommit så här långt i texten har kunnat finna någon behållning av – utan representerar även slutet på en resa som påbörjades för drygt fyra och ett halvt år sedan. Man brukar som bekant säga att det är resan som är målet, vilket öppnar upp för frågan vad denna text då skulle vara... Under åtskilliga timmar tillbringade vid mitt skrivbord medan såväl snö som regn, och en och annan gång även solsken, har pockat på uppmärksamhet från världen utanför har detta bland mycket annat rört sig i mina tankar. Säkert är dock att denna resa har kantats av fantastiska människor. Dessa har gjort till och med uppförsbackarna minnesvärda, men har också funnits där för att dela solskenet vid triumfernas ögonblick.

För tillkomsten av denna text, som kan sägas representera den sista backen innan man får blicka ut över ett nytt landskap som lockar till nya vandringar, har vissa personer haft särkilt stor betydelse. Först och främst önskar jag rikta ett varmt tack till min handledare **Harald** på KTH som har funnits till i vått och torrt och hjälpt mig bana väg när det har känts snårigt. Stort tack även till **Krister Svanberg** för all hjälp med optimeringsproblemen som har kantat denna väg. Även **Gunnar Englund** har ställt upp med vägledning när frågor kring resampling har dykt upp. Ett stort tack även till mina handledare på SEB, **Anders Nordlander** och **Jonas Ljungqvist**, som gjorde detta examensarbete möjligt, men även till **alla andra på SEB** som har ställt upp när jag har sökt svar på frågor och gjort arbetsmiljön trivsam. Tack allesammans, ingen nämnd och – definitivt – ingen glömd. Tack även alla kära **vänner och bekanta** som ställt upp när jag har behövt skingra tankarna för att kunna tänka klarare.

Avslutningsvis önskar jag tacka **min familj** som alltid har ställt upp efter alltför långa dagar följda av alltför korta nätter och som har stöttat mig i mitt sökande efter balans mellan mental stimulans och arbete med händerna. Tack till **personen som ger mig pannkakor** till frukost och låter mig få råröra sylten, trots att bären blir malträterade. Sist men inte minst vill jag **återgälda det hedersomnämnde** som jag fick i en avhandling häromåret och passa på att tacka för att jag i alla fall fick det första – och enda – med guldpenna numrerade exemplaret. Dubbel-Pet! ;)

REFERENCES

Andrews, D.W.K., Buchinsky, M., 1997. On the number of bootstrap repetitions for bootstrap standard errors, confidence intervals, and tests. Cowles Foundation Paper 1141R.

Araten, M., Jacobs, M. Jr., Varshney, P., Pellegrino, C. R.; An Internal Ratings Migration Study; The RMA Journal April 2004

Bangia, A., Diebold, F.X., Kronimus, A., Schagen, C., Schuermann, T.; Ratings migration and the business cycle, with application to credit portfolio stress testing; Journal of Banking & Finance 26 (2002) 445–474

Banachewicz, K., van der Vaart, A., Lucas, A.; Modeling Portfolio Defaults using Hidden Markov Models with Covariates; Tinbergen Institute Discussion Paper, TI 2006-094/2

Czado, C., Pflüger, C.; Modeling dependencies between rating categories and their effects on prediction in a credit risk portfolio; Appl. Stochastic Models Bus. Ind. 2008; 24:237–259

Cox, D.R., Miller, H.D., 1965. The Theory of Stochastic Processes. Chapman & Hall, London

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

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

Feng, D., Gourieroux, C., Jasiak, J.; The ordered qualitative model for credit rating transitions; Journal of Empirical Finance 15 (2008) 111–130

Fledelius, P., Lando, D., Perch Nielson, J.; Non-Parametric Analysis of Rating Transition and Default Data; Journal of Investment Management, Vol. 2, No. 2, (Q2 2004), pp. 71-85.

Frydman, H., Schuermann, T.; Credit rating dynamics and Markov mixture models; Journal of Banking & Finance 32 (2008) 1062–1075

Gagliardini, P., Gourieroux, C.; Stochastic Migration Models with Application to Corporate Risk; Journal of Financial Econometrics; Journal of Financial Econometrics, 2005: 3(2) 188

Hanson, S., Schuermann, T.; Confidence intervals for probabilities of default; Journal of Banking & Finance 30 (2006) 2281–2301

Jafry, Y., Schuermann, T.; Measurement, estimation and comparison of credit migration matrices; Journal of Banking & Finance 28 (2004) 2603–2639

Kadam, A., Lenk, P.; Bayesian Inference for Issuer Heterogeneity in Credit Ratings Migration, Working Paper Cass Business School, City University, London & Ross Business School, University of Michigan

Kiefer, N. M., Larson, C. E.; A simulation estimator for testing the time homogeneity of credit rating transitions; Journal of Empirical Finance 14 (2007) 818–835

Koopman, S. J., Lucasb, A., Monteiro, A.; The multi-state latent factor intensity model for credit rating transitions; Journal of Econometrics 142 (2008) 399–424

Lando, D., Skødeberg, T. M.; Analyzing rating transitions and rating drift with continuous observations; Journal of Banking & Finance 26 (2002) 423–444

McNeil, A. J., Wendin, J. P.; Bayesian inference for generalized linear mixed models of portfolio credit risk; Journal of Empirical Finance 14 (2007) 131–149

Mählmann, T.; Estimation of rating class transition probabilities with incomplete data; Journal of Banking & Finance 30 (2006) 3235–3256

Nickell, P., Perraudin, W., Varotto, S.; Stability of rating transitions; Journal of Banking & Finance 24 (2000) 203±227

Parnes, D.; Time series patterns in credit ratings; Finance Research Letters 4 (2007) 217–226

Rösch, D.; An empirical comparison of default risk forecasts from alternative credit rating philosophies; International Journal of Forecasting 21 (2005) 37– 51

Trück, S.; Forecasting Credit Migration Matrices with Business Cycle Effects - A Model Comparison; European Journal of Finance, 2008: 14(5) 359

Weissbach,R., Dette, H.; Kolmogorov–Smirnov-type testing for the partial homogeneity of Markov processes – With application to credit risk; Appl. Stochastic Models Bus. Ind., 2007; 23:223–234