Can macro variables improve transition matrix forecasting?

Georges Mansourati & Anders Olsson

May 2006

Abstract

Transition matrices that describe the probability for firms to migrate between rating classes are important inputs in many credit risk applications. In this thesis we try to find a simple way to forecast transition matrices that reflect actual future probabilities of credit migration better than unconditional transition matrices. By conditioning on the state of the economy measured in terms of GDP-growth we find that the state of the economy has an effect on transition matrices. We use this information together with macroeconomic indicators to produce transition matrices that are conditioned on the state of the economy. We find that the simple method we present produces better transition matrix estimates than the ordinary unconditional transition matrix.

Acknowledgements

We would like to thank our supervisor, Joel Reneby, for valuable feedback, comments and discussions. We would also like to thank Hilda Hellgren at Standard & Poor's for giving us access to, and helping us understand, the data used in the thesis. Last but certainly not least We would like to thank our families, friends, and girlfriends for their continuous support and for bearing with us during times when all of our attention was consumed by writing this master's thesis.

Contents

1 Introduction

For banks and other financial institutions it is vital to quantify the exposure to different kinds of risk. Credit risk, i.e. the risk of a debtor defaulting on his or her contractual obligations, is the biggest risk that a bank is exposed to. In recent years credit risk related issues have been actualized through the introduction of the new capital adequacy framework, commonly known as Basel II.

As a response to the growing interest for credit risk modelling, a number of commercial products have been developed for this purpose. Examples of such commercial products are JP Morgan's CreditMetrics, KMV's Portfolio Manager, Credit Suisse Financial Products' CreditRisk+ and McKinsey & Company's CreditPortfolioView. These models are often classified into one of the three following model classes. (1) Structural models, e.g. Credit Metrics and Portfolio Manager, are based on Merton's [11] model in which firms default when the value of the assets falls below the value of the liabilities. (2) Actuarial models, e.g. CreditRisk+, make use of mathematical techniques common in loss distribution modelling in the insurance industry. (3) Econometric models, e.g. CreditPortfolioView, use regression techniques to model default risk by using both macroeconomic and idiosyncratic variables (see Wilson [18] & [19]).

In a credit risk management context credit migration or transition matrices [1] are very important inputs. For example, the transition matrix is an important input in the CreditMetrics model to produce probability distributions of future portfolio values. Furthermore, from a Basel II or general risk management perspective, the transition matrix can be used to assess the riskyness of a portfolio in order to set aside a suitable amount of safety capital. In other words, having an accurately estimated transition matrix is very important.

In Jafry & Schuermann [8], two different approaches for estimation of transition matrices are evaluated and compared: the duration approach and the cohort approach. They conclude that duration approach methods produce more efficient results in terms of the regulatory capital levels they imply.

In Nickell et. al. [13] they study industry and domicile effects as well as business cycle effects on the transition matrix. They find that both business cycle and industry factors affect transition matrices.

Bangia et. al [3] also show that there is a difference between transition matrices in macroeconomic expansion and contraction. By using this knowledge, they also show that the Value-at-Risk (VaR) level for a bond portfolio will differ greatly depending on if one uses expansion or contraction matrices

[1]The two expressions credit migration matrix and transition matrix will be used interchangeably throughout this thesis.

for the VaR calculation.

To our knowledge the ability to forecast transition probabilities by using business cycle indicators has not yet been treated in the literature. I.e. we have seen that the state of the economy does have an effect on rating migration probabilities and that this effect is both statistically and economically significant. However, we have not seen how this fact to improve the prediction of the future transition matrix. Thus, the purpose of our study is to investigate whether it is possible to improve forecasts of the credit migration matrix by using business cycle indicators.

To achieve this we collect S&P data on corporate ratings from 1986 to 2005 to produce one-year transition matrices as well as an aggregate one-year unconditional transition matrix for the whole sample using the cohort approach. The reason for choosing the cohort approach instead of the duration approach is mainly due to the relative technical simplicity of the cohort method. Next, we condition transition matrices on the state of the economy, which is measured in terms of GDP-growth, to produce an aggregate conditional transition matrix for expansion and contraction periods respectively. We show that the difference between the expansion and contraction matrix is statistically significant, which is in line with previous research as mentioned above. Then we introduce two macroeconomic indicators to help us forecast the future state of the economy one year ahead.

In order to evaluate how similar two transition matrices are, we use the Euclidean L2 measure presented in Jafry & Schuermann [8]. However, the properties of this measure are not very well suited for comparison of transition matrices due to differences in the uncertainty of entries in the matrix. Therefore we present an alternative, weighted L2 metric, which accounts for such differences in the uncertainty.

Finally, by using the macroeconomic indicators and the conditional transition matrices, we show that it is possible to produce better transition matrix forecasts in terms of different L2 metrics that we introduce.

The structure of this thesis is as follows: In Section 2 we describe the dataset we have gathered from S&P and how we manage the data to fit the purpose of our study. In Section 3 we construct transition matrices and provide some preliminary analysis of the data in our study. In Section 4 we describe the theory that we base our study on. In Section 5 we develop and explain our forecasting methodology. Finally, in Section 6 we present and discuss our results.

2 Data

In this section we describe the data that we gathered at S&P in Stockholm. We explain how we transform the data set to fit the purpose of our study and we give some descriptive statistics for the set. We conclude this section by providing a brief discussion on potential problems associated with the kind of data we use.

Most previous research treating credit risk in general (and analysing the transition matrix in particular) is based on data from S&P or Moody's, the two leading credit rating agencies. Using any of the two seems to be suitable for the purpose of our study since they provide the highest number of credit ratings. Furthermore, using data from one of these two agencies simplifies comparison to the results of other studies. We choose to use ratings from S&P in our study.

The dataset we use consists of American issuer rating migrations from January 1, 1986 to September 30, 2005. Issuer rating migrations are based on harmonised unsecured bond data with a specified long term maturity. M&A and group effects are also taken into account in this kind of rating. We focus on American companies because of the larger number of rated companies in the US compared to Europe and the rest of the world (see Bangia [3]). Performing a study on American data also simplifies comparison to other studies.

Our original data consists of three columns. The first column is the name of the rated company. The second column is the date when the company received its first rating or the date when the company's rating was changed. The third column is the rating that the company was assigned on the given date (i.e. the new rating). Note that in this dataset each company potentially appears at several rows. We use these three columns to construct time series for each company, showing its credit rating at discrete quarterly points in time. In Section 3 we describe how these series are used to create transition matrices which we use in our analysis.

Each company is assigned a rating according to S&P's scale ranging from AAA which is the highest, down to C which is the lowest rating. The complete set of categories is AAA, AA, A, BBB, BB, B, CCC, CC, C and D which symbolizes default. The default rating is used when issuer payments are not made on the date due, unless Standard & Poor's believes that such payments will be made during the grace period. The D-rating is also used upon the filing of bankruptcy or a similar action jeopardizing the issuer's payments. There are also so called modifiers in certain rating classes. E.g. AA+, AA and AA- are all separate ratings within the AA rating such that AA+ is a higher rating than AA (which is higher than AA-). In addition to these categories there is a special rating category, Not Rated (NR), which is assigned to a company if its rating is withdrawn.

The S&P credit ratings are based on three main criteria: an analysis of the industry in which the firm operates, a qualitative judgement of the company and a numerical analysis of the firm's historical and future performance, both financially and operationally. Companies are rated on the request of the company itself or on the request of another interested institution. For further information on the rating procedure we refer to Standard

& Poor's [15], and Altman et. al. [2].

To conclude this section we want to point out that there are potential problems with the kind of data set we use. Essentially we are basing our analysis on data which is the subjective judgement of a third party (S&P) determining how credit worthy a company is at a point in time. One problem could be a change in the routines and criteria that rating agencies use when assigning credit ratings. This may occur due to conflicts of interest. E.g. certain mutual funds are prohibited by law to invest in non-investment grade corporations .[2] In this situation it is desirable for the company to attain a higher rating (e.g. to be able to get less expensive funding). The mutual funds wish to have a larger set of companies to choose from when investing. Finally, the rating agency wants to get paid for providing companies with a rating. With this in mind, companies which are likely to receive a lower rating might prefer to withdraw their rating, i.e. stop paying for being rated. This could lead credit rating institutions to act short-sightedly and lower their requirements. For a more detailed discussion on these issues we refer to Altman et. al. [2].

It is also worth mentioning that the number of rated companies has increased drastically in recent years. Figure 1 shows the number of rated companies in the beginning of each year in our sample period. With this in mind we should note that data from more recent years will have a bigger impact on our estimates than data from earlier years.

3 Properties of transition matrices

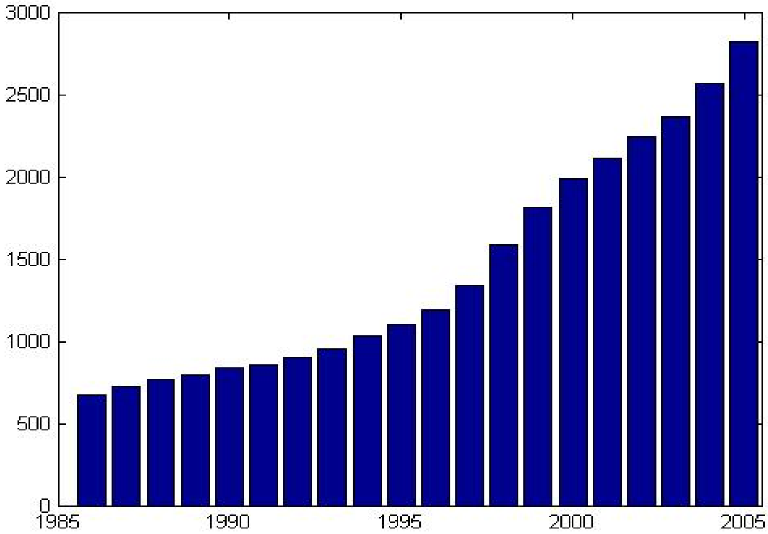
A transition matrix displays the historical probability of rating migrations. In this section we describe how the transition matrix is calculated and we also study the stability of transition matrices over time for different transition horizons.

3.1 Creating the transition matrix

We will use what is referred to in the literature as the cohort approach when constructing the transition matrix. This is the method of choice for most practitioners (c.f. Jafry & Schuermann [8]) and to our knowledge, the most simple and intuitive way to calculate the transition matrix. Using this approach, the elements of the migration matrix, i.e. the probability that a company migrates between two rating classes in period t, are calculated as $p_{ij}(t) = \frac{\#(i \to j)}{\sum_{m=1}^{9} \#(i \to m)}$. In other words, an element (i, j) of the transition matrix reflects the historical frequency of migrations from rating i to rating j. More formally:

[2]The classes AAA to BBB are regarded as investment grade ratings and BB to CCC are non-investment grade.

Figure 1: *The number of rated companies in the beginning of each year in our sample period.*



$$p_{ij}^{t} = \frac{\sum_{k=1}^{n_t} I_k(i \to j)}{\sum_{k=1}^{n_t} \sum_{m=1}^{9} I_k(i \to m)} \tag{1}$$

where i stands for the rating class in the beginning of the period, j stands for rating class in the end of the period, 9 is the number of rating classes and n_t is the number of rated companies in the beginning of period t. State 1, i.e. $i = 1$, refers to the highest rating, AAA, and state $i = 7$, referring to CCC-C, is the worst. State 8 and 9 correspond to default (D) and Not Rated (NR) respectively. I_k is the indicator function for company k defined as:

$$I_k(i \to j) = \begin{cases} 1 & \text{if company } k \text{ migrates from rating } i \text{ to rating } j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

In other words, the numerator in (1) is the number of companies changing rating from i to j in period t. The denominator is the number of companies changing rating from i to any rating $j = 1, \ldots, 9$ in period t.[3]

In (1) we describe the calculation of the actual *one-period* transition matrix, $M(t)$, in period t. In (3) we calculate an aggregate transition matrix for the whole sample period. Again, using the cohort approach this matrix is calculated as:

[3]Put differently, the denominator is the number of companies starting out in rating i.

$$p_{ij}(T) = \frac{\sum_{t=1}^{T}\sum_{k=1}^{n_t} I_k(i \to j)}{\sum_{t=1}^{T}\sum_{k=1}^{n_t}\sum_{m=1}^{9} I_k(i \to m)} \tag{3}$$

where T is the total number of transition periods in our sample period. In this case the numerator is the aggregated number of observed rating migrations from i to j during the transition periods. The denominator is the aggregated number of observed rating migrations changing rating from i to any rating $j = 1, \ldots, 9$ (i.e. the number of observations starting out in i).

The probability of migrating to rating class D (the probability of default, PD) is not calculated in the same way as the other migration probabilities. The probability of migrating between two non-default rating classes is simply calculated by observing ratings at two separate discrete points in time. However, the probability of migrating to D is defined as the probability of getting a D rating at least once during the transition period in question. This is what is referred to in the reference literature as an absorbing default state (c.f. Jafry & Schuermann [8]). It is referred to as absorbing since a company that has once entered the default state during a certain transition period, stays in the D rating class until the end of the period in order to enter the numerator in the calculation of the PD quotient. This special treatment of the PD is intuitive since this probability should reflect the fact that a company defaults some time during the transition period in question, not only that the company is in the state of default at the end of the period.

By using the total number of rated companies in each time period t we ensure that our sample grows and has relevant business industry composure.[4] This procedure is associated with advantages and disadvantages. The advantage is that we get more reliable estimates by adding more companies to our sample as they are rated by S&P. One disadvantage is that we are including businesses from many years back that might have had different migration behaviour e.g. due to differences in regulation etc.. Our approach assigns each company credit migration observation an equal weighting, thus more recent years receive a higher weight since more rating migrations are observed.

There are also disadvantages associated with the use of a cohort approach when studying transition matrices. One such disadvantage is that the cohort approach only takes one migration during the transition period into consideration. For example if a company migrates first from rating class AAA to A and then to BBB, all during one year, the only one-year transition probability that is affected is the probability of migrating from AAA to BBB.

[4] As opposed to sticking to an initial sample of companies which would get smaller over time as some companies default. Moreover, by sticking to an initial sample of companies we would not have a relevant industry composure in the sample (e.g. we would perhaps leave out IT-companies that emerged in the 90's).

An alternative way of creating the transition matrix is suggested in Jafry & Schuermann [8]. This method is called the duration approach as it takes into consideration the time spent in each rating class within the sample period. Put differently, this method uses all transitions within the sample, and the time spent in each rating class when constructing the transition matrix. However, we believe that the duration approach is unnecessarily complicated for the purpose of our study. Furthermore, the cohort approach has been more frequently used in studies prior to ours which simplifies comparison of our results.

On the main diagonal in the transition matrix the number of observations is large. The further away from the main diagonal, i.e. the more extreme rating movements, the fewer are the observations. This fact is also reflected in the standard errors of the elements in the transition matrix. In other words, elements further away from the main diagonal will have larger standard errors.[5] To mitigate this problem we remove the modifiers to get rating categories with more observations in each rating class. E.g. A-, A and A+ are all considered as A ratings. This procedure is common practice (c.f. Bangia [3] and Nickell et. al. [13]). In our case this results in $N = 9$ rating classes including D and NR and considering CCC, CC and C as one group, denoted CCC-C.

Another interesting issue that arises when constructing transition matrices is how to handle withdrawn ratings. Some studies treat migration to the 'Not Rated', NR, category as a negative migration indicating that the company in question withdrew its rating before potentially being down rated. Other studies treat migration to NR as non-informative (c.f. Bangia [3]). The transition to NR is interesting in the light of the new Basel II accord. When calculating the risk-weighted capital under the standard approach for credit risk the risk-weight is smaller for NR companies (100%) than for non-investment grade (150%) [6] companies. Thus, for banks it is more profitable to include NR companies in their lending portfolio than non-investment grade companies. This fact could potentially lead companies which are in the danger zone for downgrade to withdraw their rating. In the long run this might lead rating agencies to soften the requirements for investment grade ratings. In this thesis we will not analyze migration to NR but only use it as any other rating class.

To give a concrete example of what a transition matrix looks like we present the aggregated one-year unconditional transition matrix (UTM) in Table 1. The rows in the matrix correspond to the initial rating and the columns correspond to the final rating one year later.

We can observe some important features of the transition matrix. First we should note that the entries in each row sum to unity. We also note that

[5] See matrices with standard deviations and coefficients of variation in Appendix A.

[6] S&P ratings BB or lower are regarded as non-investment grade (or junk bond grade).

Table 1: *The yearly Unconditional Transition Matrix for the period: January 1, 1986 - December 31, 2004.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**94.4**	4.6	0.6	0.0	0.1	0.0	0.0	0.0	0.2
AA	0.5	**88.8**	9.8	0.6	0.0	0.0	0.0	0.0	0.3
A	0.0	1.7	**91.4**	5.8	0.2	0.1	0.1	0.0	0.7
BBB	0.0	0.1	4.1	**89.7**	4.0	0.6	0.1	0.0	1.3
BB	0.0	0.1	0.1	7.2	**83.3**	7.2	0.5	0.1	1.5
B	0.0	0.0	0.0	0.4	8.9	**85.0**	3.1	0.7	1.8
$CCC-C$	0.6	0.0	0.0	0.0	3.5	21.1	**56.6**	7.2	11.0

the highest probabilities are found on the main diagonal. This corresponds to the probability of a company staying in its initial rating class. Another important feature is that lower rating classes exhibit a lower probability of remaining in the same rating class compared to higher ratings, i.e. rating movements are more frequent for low rating classes. Furthermore, we note that the probability of default is higher for lower rating classes. Finally, we note that the entries in the matrix are typically not monotone on all rows in the sense that probabilities on the same row, do not decrease monotonically with the distance from the main diagonal element (c.f. Jafry & Schuermann [8]). For example we observe that a CCC-C rated company is more likely to migrate to AAA than to AA (or A).

3.2 Determination of transition period

In this subsection we study the stability of the transition matrix over time. This will help us determine which transition period are reasonable to use for the purpose of our study.

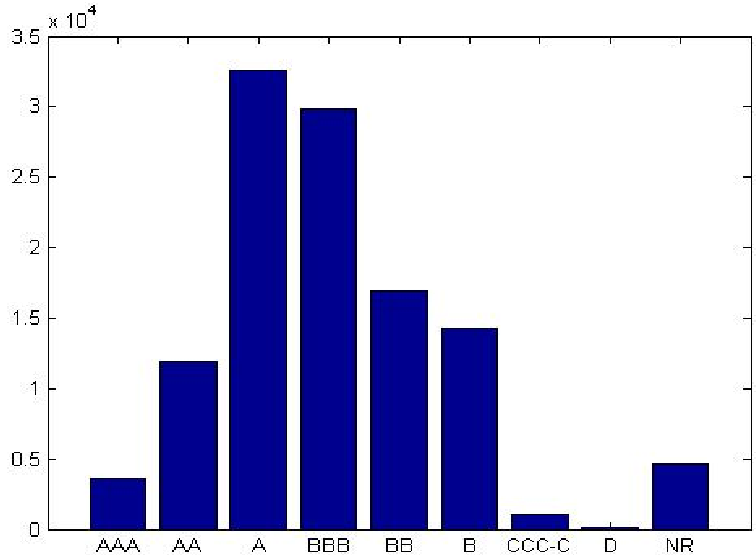
By stability we mean how the elements of the transition matrix, M_t, vary over time. We choose to study the four elements on the main diagonal for which we have the highest number of observations, since we expect these estimates to be the most stable over time.

If these elements are unstable, i.e. fluctuate too much over time, we can expect the other elements in the same transition matrix to be even more unstable due to fewer observations. To be able to study the effect of macro variables on the transition matrix we need to observe statistically significant variations in the transition probabilities over time.

Naturally we expect longer transition periods to yield more stable transition probabilities.[7] However, there are drawbacks associated with using too

[7]This is in line with Jafry & Schuermann [8], where they observe less noise for longer

Figure 2: *The aggregate number of companies in each rating class in the beginning of each quarter in our sample.*



long transition periods. First, the number of observed transition matrices will decrease the longer the transition period. Second, the lost information within the transition period increases since all rating changes within the sample are not taken into account.

In Figures 3 and 4 we present the development of the four mentioned elements over time, for quarterly and yearly transition periods. We see that the elements fluctuate violently for the quarterly transition period and we see no pattern that is likely to be forecasted by some macro variable. The second figure shows the same elements over time for a yearly transition period. In this case we see that the elements vary more smoothly over time.

4 Macro dependence modelling

In this section we explain why we use GDP-growth as a conditioning macro variable. We also describe how we construct conditional transition matrices, High (H) and Low (L), which depend on the magnitude of GDP-growth.

In the reference literature (c.f. Nickell et. al. [13] and Bangia [3]) different business cycle indicators are used to proxy for the state of the economy. In general, these indicators rely heavily on GDP-growth. Other reasons for conditioning on GDP-growth is the availability of data and the

transition matrices.

Figure 3: *The four elements on the main diagonal of the quarterly transition matrix corresponding to* $i = j = 3, 4, 5, 6$ *in each time period.*

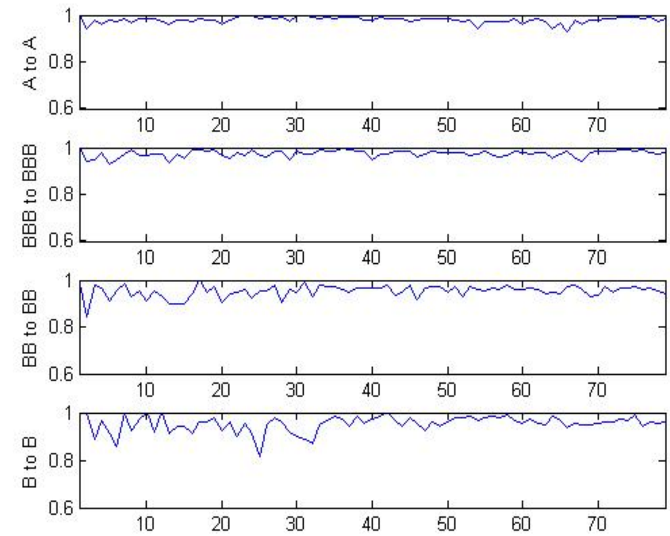
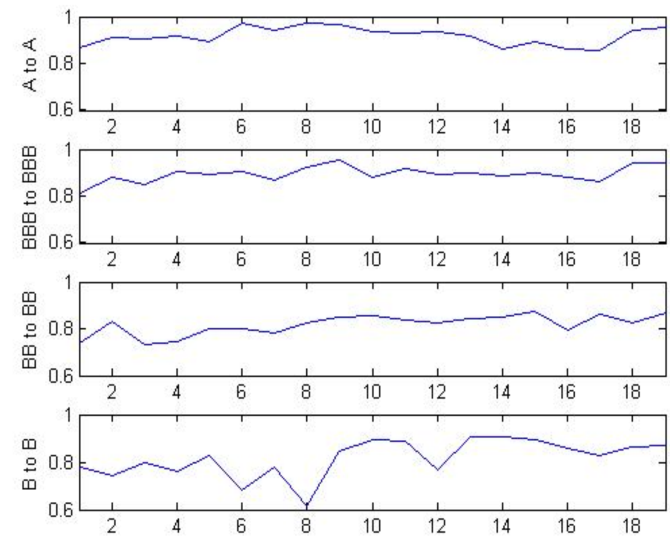


Figure 4: *The four elements on the main diagonal of the yearly transition matrix corresponding to* $i = j = 3, 4, 5, 6$ *in each time period.*



knowledge of what the variable stands for.

During the start-up phase of this study we also considered conditioning on inflation, interest rates or the NBER indicator. However, the preliminary results from using these variables as proxies for the state of the economy weren't satisfying. Moreover, since previous research has already shown that business cycle indicators that heavily rely on GDP-growth have an effect on migration probabilities, we choose to use GDP-growth directly as a proxy for the state of the economy.

Below we describe how the conditioning on macro variables is actually realized, i.e. how the Unconditional Transition Matrix (UTM) and the conditional, H- and L matrices, are created. As mentioned in Section 3 the elements of UTM are calculated according to equation (3):

$$p_{ij}(T) = \frac{\sum_{t=1}^{T}\sum_{k=1}^{n_t} I_k(i \to j)}{\sum_{t=1}^{T}\sum_{k=1}^{n_t}\sum_{m=1}^{9} I_k(i \to m)}$$

I.e. this matrix is simply the aggregate transition matrix for the entire period. This is typically the matrix we use when we have no other information on the behaviour of companies in good or bad states of the economy (or any other relevant information for that matter).

On the other hand, the conditional transition matrix takes into account some macroeconomic variable. In our case the variable we condition on is GDP-growth. The formula we use to calculate the H matrix, i.e. the matrix corresponding to high GDP-growth values, is given in (4).

$$p_{ij}^{H}(T) = \frac{\sum_{t=1}^{T} I_t^H(GDP_t > GDP^*)\Big(\sum_{k=1}^{n_t} I_k(i \to j)\Big)}{\sum_{t=1}^{T} I_t^H(GDP_t > GDP^*)\Big(\sum_{k=1}^{n_t}\sum_{m=1}^{9} I_k(i \to m)\Big)} \tag{4}$$

where I_t^H, the conditioning indicator, is defined as:

$$I_t^H(GDP_t > GDP^*) = \begin{cases} 1 & \text{if } GDP_t > GDP^* \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Naturally the corresponding formula applies for the L-matrix with the adjustment that the conditioning indicator is $I_t^L(GDP_t \leq GDP^*)$.

Here, GDP^* is a threshold value that we choose ourselves. The value of GDP^* is the value that separates between good and bad states of the economy in our model. A good state where $GDP_t > GDP^*$ is labelled 'H' (for high) and a bad state, where $GDP_t \leq GDP^*$, is labelled 'L' (for low). The choice of GDP^* is not obvious and it depends to a large extent on what one wishes to study. In our case, we want to study differences in the transition matrix in periods of expansion and contraction respectively. For this purpose we believe it is reasonable to set GDP^* equal to potential growth since this is in some sense the long term average real growth of the economy. However, there is no obvious value for potential growth either.[8] Therefore we simply set GDP^* equal to the average GDP-growth over the period we are dealing with. This can be seen as a proxy for potential growth over the period we study. Furthermore, this choice insures that we get approximately the same number of H- and L periods which is desirable from a statistical point of view.

In Tables 2 and 3 we present the resulting H and L matrices when we condition on GDP-growth over the whole sample period as described above.

Table 2: *The H matrix, i.e. the transition matrix in expansion.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**93.3**	5.3	0.9	0.0	0.2	0.0	0.0	0.0	0.2
AA	0.8	**90.9**	6.8	0.8	0.1	0.1	0.0	0.0	0.6
A	0.0	1.7	**92.3**	5.0	0.1	0.1	0.0	0.0	0.8
BBB	0.0	0.1	3.9	**90.9**	3.7	0.4	0.0	0.0	1.2
BB	0.0	0.0	0.1	7.3	**83.4**	6.7	0.2	0.1	2.1
B	0.0	0.0	0.0	0.4	9.3	**86.4**	2.1	0.3	1.5
$CCC-C$	0.0	0.0	0.0	0.0	3.0	19.5	**62.0**	4.5	11.0

Table 3: *The L matrix, i.e. the transition matrix in contraction.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**95.7**	3.8	0.3	0.0	0.0	0.0	0.0	0.0	0.3
AA	0.1	**86.7**	12.8	0.4	0.0	0.0	0.0	0.0	0.1
A	0.1	1.7	**90.4**	6.8	0.4	0.1	0.1	0.1	0.5
BBB	0.0	0.2	4.3	**88.4**	4.4	1.0	0.2	0.1	1.4
BB	0.0	0.1	0.1	7.0	**83.2**	7.7	1.0	0.3	0.6
B	0.0	0.0	0.0	0.4	8.5	**83.5**	4.3	1.3	2.0
$CCC-C$	1.4	0.0	0.0	0.0	4.1	23.3	**49.3**	11.0	11.0

We note that (in general) the probabilities of downgrade, i.e. elements above the main diagonal, are lower for the H matrix than for the L matrix (and vice versa, i.e. probabilities of upgrade are higher for the H matrix than for L-matrix.) just as we would expect. The interesting question however, is

[8] One can express a belief about historical potential growth by observing a trend over a certain historical time period. However, historical potential growth need not be the same as present potential growth.

to find out if the difference between two corresponding estimated transition probabilities, $\hat{p}_{ij}^H(T) - \hat{p}_{ij}^L(T)$, is statistically significant (and of the expected sign).

In order to study whether or not the difference is significant, we make use of the fact that the number of transitions from rating class i to rating class j is binomially distributed, $n_{ij} \sim Bin(n_{i\bullet}, p_{ij})$, where $n_{i\bullet} = \sum_{j=1}^{9} n_{ij}$ and p_{ij} is the *true* probability to go from rating i to j. This approach, which is presented in Nickell et. al. [13], greatly simplifies our inference problem.

Using the binomiality assumption, the inference problem formulated above turns out to be a common problem in statistics to which there is a straight forward solution (see Blom [4]).

Table 4: *The Difference between entries in the H matrix and the L matrix, i.e.* $p_{ij}^H - p_{ij}^L$. * *stands for statistically significant difference at a 5% significance level.*

10^{-2}	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**−2.3**	1.5	0.7	0.0	0.2	0.0	0.0	0.0	0.0
AA	0.7	**4.2***	−5.9*	0.4	0.1	0.1	0.0	0.0	0.5
A	0.0	0.0	**1.9***	−1.8*	−0.3	0.0	−0.1	−0.1	0.4*
BBB	0.0	−0.1	−0.4	**2.5***	−0.7	−0.6*	−0.2	−0.1	−0.3
BB	0.0	−0.1	0.1	0.3	**0.2**	−1.0	−0.8	−0.2	1.5
B	0.0	0.0	0.0	0.0	0.8	**2.8***	−2.2*	−1.0	−0.5*
$CCC-C$	−1.4	0.0	0.0	0.0	−1.1	−3.8	**12.7**	−6.5	0.0

In Table 4 we see that there are only a few elements, concentrated around the main diagonal, that are significantly different from zero on a 5% significance level. Moreover, the significant differences are all found above the main diagonal. This proves that companies are more frequently up-rated in expansion periods than in contraction periods.

An element in position (i, j) is said to be significantly different from zero if it fulfills the two following criteria:

1. $\hat{p}_{ij}^H(1-\hat{p}_{ij}^H)n_{i\bullet}^H > 10$ and $\hat{p}_{ij}^L(1-\hat{p}_{ij}^L)n_{i\bullet}^L > 10$

2. $\left| \frac{\hat{p}_{ij}^H - \hat{p}_{ij}^L}{\sqrt{\frac{\hat{p}_{ij}^H(1-\hat{p}_{ij}^H)}{n_{i\bullet}^H} + \frac{\hat{p}_{ij}^L(1-\hat{p}_{ij}^L)}{n_{i\bullet}^L}}} \right| > \lambda_{0.025} = 1.96$

The first criterion is to insure that the normal approximation for the binomially distributed stochastic variables $n_{i\bullet}^H$ and $n_{i\bullet}^L$ respectively is good enough. The second criterion is the actual 5% significance test.

By applying this procedure we are assuming that the *number* of transitions, n_{ij}, follows a constant parameter multinomial process. This means that we are indirectly assuming that the rating process for each company over time is a Markov process i.e. that the rating of a company in time t only depends on the rating in time $t-1$. This is a very practical assumption to make since it greatly simplifies our study. However, it has been shown in several papers that this is not the case for ratings over time. In Bangia [3] for example, they show that there is an up- and downward drift effect in ratings. E.g. a company which has received a higher rating between time $t-2$ and $t-1$ has a higher probability of being up-rated again between $t-1$ and t compared to a company which hasn't been up-rated previously. However, we argue that this effect is not too serious for the purpose of our study.

5 Forecasting the transition matrix

In Section 4, Table 4, we have seen that conditioning on GDP-growth generates statistically significant, and intuitively correct differences between the elements of the H and L matrices. In this section we suggest a simple way to make use of this fact when forecasting future transition matrices. We present indicators that will be used to forecast the future state of the economy (expansion or contraction). Finally, to evaluate how similar two transition matrices are, we introduce some distance metrics.

5.1 Forecasting the future state

In this thesis we ask ourselves if we can use simple methods to produce better forecasts of the future transition matrix than just using the historical unconditional transition matrix (UTM). Therefore we choose a methodology where forecasting of the future state of the economy is done exogenously in some sense.[9] By future state of the economy we mean: will the economy experience higher or lower than potential GDP-growth one year from now?. I.e. we wish to decide in which of two states, H or L, the economy will be in one year. Expansion periods which experience high growth are labelled H and contraction periods which experience low growth are labelled L.

To be able to distinguish future expansion- and contraction periods we use two different external indicators which have proven to be useful for this purpose.

The first indicator is the slope of the yield curve estimated by the difference between a long and a short yield. This indicator is suggested by Estrella

[9]Obviously this approach would have been best suited if we could produce our own GDP forecasts using for example Vector Autoregressive or Dynamic Stochastic Genereal Equilibrium models. However, producing such forecasts is not trivial and taking on such an approach would be very time consuming.

& Hardouvelis [7] to have predictive power of real economic activity. We follow their definition and calculate this indicator as:

$$SPREAD_t = R_t^{LONG} - R_t^{SHORT} \tag{6}$$

where R^{LONG} is the 10-year government bond rate and R^{SHORT} is the 3-month T-bill rate. Both yields are annualized bond equivalent yields.

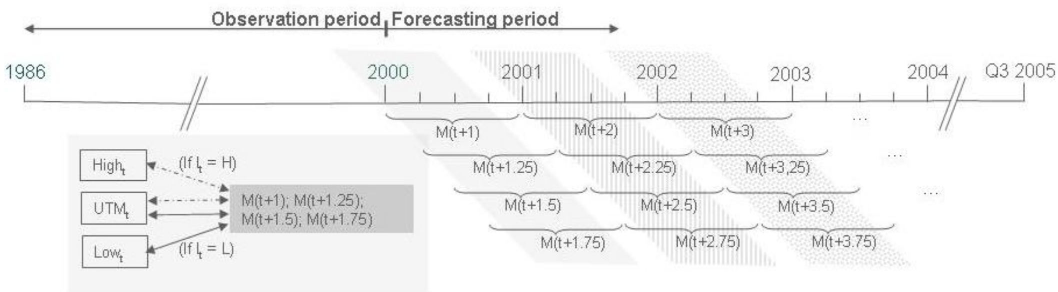
The second indicator we chose is the Purchasing Manufacturing Index (PMI) which is published by the Institute for Supply Management. PMI is a weighted index based on new orders (30 %), production (25 %), employment (20 %), supplier deliveries (15 %) and inventories (10 %). This index takes values between 0 and 100 where values above 50 indicate that the manufacturing economy is generally in expansion (and vice versa).[10] The PMI has a track record of anticipating economic growth according to different studies. We therefore argue that PMI can be used as an indicator for future economic growth in our study.

To have a benchmark, we also use actual GDP-growth for the future period as a third indicator. I.e. by using this indicator we are assuming that we can predict the future state of the economy perfectly.

5.2 Comparing transition matrices

We start by describing how we use the conditional transition matrices, H and L matrix defined in Section 3, to forecast future transition matrices. The objective is to see whether these matrices can be used to obtain transition matrices that better reflect future rating migrations than just using the unconditional historical transition matrix which disregards from business cycle effects. We present the methodology we apply graphically in Figure 5.

Figure 5: *Graphical representation of the forecasting evaluation.*



[10] www.ism.ws

To estimate the first four transition matrices $M(2000:1)$, $M(2000:2)$, $M(2000:3)$ and $M(2000:4)$, the sample on which we estimate an H-, an L- and a UTM matrix consists of all rating transitions between January 1986 and December 1999. The remainder of our dataset is used for out-of-sample forecasting. To estimate $M(2001:1)$, $M(2001:2)$, $M(2001:3)$ and $M(2001:4)$ the observation period is extended by one year and consists of rating transitions from January 1986 to December 2000 etc..

Thus, each year t in the out-of-sample period corresponds to four yearly transition matrices (except the year 2004 which only corresponds to three). Each matrix $M(y:q)$ is constructed by applying the cohort approach presented in Section 3, equation (1). Consequently, $M(y:q)$ is defined as the observed transition matrix over one year, starting in year y, quarter, q and ending at year y+1, quarter q-1. For example, using this notation $M(2000:2)$ denotes the transition matrix observed for companies from the beginning of the second quarter in 2000 to the end of the first quarter in 2001. Thus, using overlapping transition matrices as described, we have 19 out-of-sample transition matrices to test our forecasting methodology on.[11]

Now that we have described how to construct the components used in the forecasting, we move on to explain how the actual forecasting is done and evaluated. The idea is rather simple. Instead of forecasting the future transition matrix itself, we try to forecast the future state of the economy. Depending on our belief about the future state of the economy we estimate the future transition matrix M with either the H or the L matrix. I.e. if our macro forecasting, using the indicators presented in the first subsection of Section 4, suggests that the future period is an expansion period, with above average GDP-growth, we will use the H matrix to proxy for the future transition matrix.

To determine which matrix of the conditional H (or L) matrix and the uncoditional transition matrix, UTM, is closest to the actual transition matrix M, we introduce the standard L2 distance metric d. This is a metric used for example in Jafry and Schuermann [8].

$$d(X,Y) = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - y_{ij})^2} \tag{7}$$

where $m = 7$ and $n = 9$ for the transition matrices that we study. The drawback of the distance metric d introduced above is that it puts equal weight on the distance between all element pairs. This is an unwanted property since there can be elements in the transition matrix far from the main diagonal that are rather volatile over time. This makes the distance metric d instable due to high uncertainty in certain estimated migration

[11] We have four matrices for each of the years 2000, 2001, 2002 and 2003. However, we only have three matrices for 2004 since our sample ends in the third quarter of 2005.

probabilities. One way to mitigate this problem is to study the distance between a limited number of elements in the transition matrix that are of more interest. For example, a set of interesting elements to study are those in the default column. I.e., by studying the distance between the default columns of two matrices X and Y and comparing it to the distance between the default columns of Z and Y, we can tell which one of X and Z has a default column closest to the default column of Y. Such a distance measure is introduced below and is denoted by d_D

$$d_D(X,Y) = \sqrt{\sum_{i=1}^{m}(x_{i,8} - y_{i,8})^2} \tag{8}$$

Both distance measures introduced above do not directly take into account the uncertainty in the estimates of the different elements of the transition matrix. They weigh the squared difference between each pair of elements equally. To account for the higher uncertainty in certain elements when measuring the distance between transition matrices, we introduce the weighted distance measures d_w and d_{Dw} below.

$$d_w(X,Y) = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{(x_{ij} - y_{ij})^2}{CV(y_{ij})}} \tag{9}$$

$$d_{Dw}(X,Y) = \sqrt{\sum_{i=1}^{m}\frac{(x_{i,8} - y_{i,8})^2}{CV(y_{i,8})}} \tag{10}$$

Again, (9) takes into account all entries in the transition matrix while (10) only considers the eighth column, i.e. the probabilities of default. The difference between these two measures and the measures introduced in (7) and (8) is the weight factor $CV(y_{ij}$ in the denominator, where CV denotes the coefficient of variation of the entry on row i column j in the Y matrix, and is defined as

$$CV(y_{ij}) = \frac{\sigma_{ij}}{\mu_{ij}} \tag{11}$$

In (11) σ_{ij} denotes the standard deviation of the migration probability p_{ij} and μ_{ij} denotes the expected value of p_{ij}. μ_{ij} in the denominator is simply replaced by the observed historical default frequency $\hat{p}_{ij}$ from (1). σ_{ij} in the numerator is easily calculated under the binomiality assumption (see Nickell et. al [13]) as

$$\sigma_{ij} = \frac{\hat{p}_{ij}(1 - \hat{p}_{ij})}{\sqrt{n_{i\bullet}}} \tag{12}$$

Note that in (9) and (10) we only use the coefficient of variation for the Y matrix. The reason for considering only the uncertainty in the second matrix (i.e. the matrix denoted Y in (7)-(10)) is because we use the distance metrics to compare for example the two distances: $d_w(UTM_t, M_{t+x})$ and $d_w(H_t, M_{t+x})$ to see which of UTM_t and H_t is closest to M_{t+x} when we have forecasted a future expansion period. In other words, in (9) we have first that $X = UTM_t$ and $Y = M_{t+x}$ and in the second case that $X = H_t$ and $Y = M_{t+x}$. Thus, by considering the coefficient of variation only for the M matrix we insure that the same weight coefficient is used in both cases, which is desirable.

6 Results and concluding discussion

In this section we present the results of our forecasting together with a concluding discussion. We also suggest some topics for future research.

The results in Table 5 show how many times (in percent) out of 19, the L or the H matrix was a better approximation to the M matrix than the UTM matrix.

Table 5: *The fraction of times (in percent, out of 19) that the L2 distance between the H or L matrix and the M matrix is smaller than the L2 distance between the UTM and the M matrix. The standard deviation under an assumption of Binomiality is presented within brackets. A '*' marks that the value is significantly different from 50 % at the 5 % level.*

	SPREAD	PMI	GDP
Threshold	1.2 (%)	50	3.0 (%)
Entire matrix (%)	63	58	63
	(11.1)	(11.3)	(11.1)
Default column (%)	68*	53	63
	(10.7)	(11.5)	(11.1)
Entire matrix, weighted (%)	63	74*	68*
	(11.1)	(10.1)	(10.7)
Default column, weighted (%)	74*	47	68*
	(10.1)	(11.5)	(10.7)

The results in Table 5 show that, on average, the L or H matrix do a better job at forecasting the transition matrix than the UTM matrix. Only in the case where we use PMI as an indicator for future real growth, does the UTM matrix do a marginally better job in forecasting M when we use the weighted L2 measure for the default column.

To test the stability of our results we perturb the threshold values for the indicators slightly. This will make certain periods switch from H to L periods and vice versa. These results are presented in Tables 12 and 13 in Appendix B.

The objective of our thesis is to provide a simple and intuitive way to improve transition matrix forecasting. Assuming that GDP-growth is known removes the macro-variable forecast effect and only indicates the predictive properties of our estimated transition matrices. Thus analysing the forecast results assuming that the future state of the economy is known gives us a better understanding of the results using the other two indicators. This case should be seen as a benchmark when evaluating the performance of the other two indicators. The forecast results assuming GDP growth is known are all above 50%. For the weighted L2 measures the results are also significant at the 5% level. By perturbing the threshold with 5 additional basis points in each direction we see that the results (see Appendix B), especially the weighted L2 measures, are better for our estimated matrices. This strengthens our belief that our forecast methodology can improve transition matrix forecasting.

It is notable for PMI that the results are very different depending on which distance measure is used. This can either suggest that PMI is not a very good indicator of future economic growth for our sample. Or it is simply a consequence of the limited number of transition matrices we compare. The reason for not choosing a longer time period for forecasting is the relatively small number of corporate ratings in the years 1985-2000 (cf. Figure 2). Moving one or more of the later years from the observation period to the forecasting period would generate inferior conditional transition matrices.

The results in Table 5 are not very surprising since we have already seen, in for example Bangia [3] and Nickell et. al. [13], that macro variables have an effect on transition matrices. What we have done is to present a simple method for how to use this information to actually produce better transition matrices in the L2 sense. We have also suggested a weighted L2 metric to account for the difference in relevance between the squared pairwise differences.

In the introduction we state that the purpose of our study is to investigate whether it is possible to improve migration matrix forecasts by using business cycle indicators. Judging from the results that are presented in Table 5 we conclude that it seems possible do so. However, it would be interesting to see if more of the entries in Table 5 would turn out to be statistically significant if we would have had a dataset from a longer time period that would enable us to use a longer forecasting period. The results presented above are based on five estimated H and L transition matrices compared to 19 different M matrices. This provides a very limited number of forecast observations, thus our results should be interpreted cautiously.

Suggestions for future research are many. For example it would be interesting to see how our method could be used to generate money in some kind of trading strategy. In this thesis we have used a method that is based on the cohort approach when creating transition matrices. Since this is not the most accurate way to produce transition matrices we suggest that the duration approach (mentioned in Section 3) is implemented instead of the cohort approach.

References

[1] Altman, E., Brooks, B., Resti, A. & Sironi, A., 2003, The Link between Default and Recovery Rates: Theory, Empirical Evidence and Implications.

[2] Altman, E., Caoette, J. & Narayanan, P. , 1998, Managing Credit Risk - The Next Great Financial Challenge, John Wiley & Sons, Inc.

[3] Bangia, A., 2002, Ratings Migration and Business Cycle, With Application to Credit Portfolio Stress Testing, Journal of Banking & Finance 26, 445-474.

[4] Blom, G., 1989, Sannolikhetsteori och statistikteori med tillämpningar, Studentlitteratur.

[5] Delianedis, G. & Geske, R., 1998, Credit Risk and Risk Neutral Default Probabilities: Information About Migration and Defaults.

[6] Duffie, D., Wang, K. & Saita, L., 2006, Multi-period corporate default prediction with stochastic covariates, NBER Working Paper 11962.

[7] Estrella, A. & Hardouvelis, G., 1991, The Term Structure as a Predictor of Real Economic Activity, The Journal of Finance, Vol. 46, No. 2.

[8] Jafry, Y. & Schuermann, T., 2004, Measurement, estimation and comparison of credit migration matrices, Journal of Banking & Finance 28, 2603-2639.

[9] Jarrow, R., Lando, D. & Turnbull, S., 1997, A Markov Model for the Term Structure of Credit Risk Spreads, The Review of Financial Studies, Vol. 10, No. 2, 481-523.

[10] McNeil, A. & Wendin, J., 2005, Dependent Credit Migrations, Working Paper Series at FINRISK, Working Paper No. 182.

[11] Merton, R., 1974, On the pricing of corporate debt: the risk structure of interest rates, The Journal of Finance, Vol. 29, No. 2.

[12] The NBER's Business-Cycle Dating Procedure.

[13] Nickell, P., Perraudin, W. & Varotto, S., 2000., Stability of rating transitions, Journal of Banking & Finance 24, 203-227.

[14] Saunders, A. & Allen, L., 2002, Credit Risk Measurement New Approaches to Value at Risk and Other Paradigms, Wiley Finance.

[15] Standard & Poor's, Rating av Företag.

[16] Standard & Poor's, Rating Definitions, www2.standardandpoors.com.

[17] Tarashev, N., 2005, An empirical evaluation of structural credit risk models, BIS Working Papers, No 179.

[18] Wilson, T., 1997, Portfolio Credit Risk (I), Risk, Vol. 10, No 9.

[19] Wilson, T., 1997, Portfolio Credit Risk (II), Risk, Vol. 10, No 10.

A Additional data

In this appendix we present matrices with the number of companies changing rating class and standard deviations.

Table 6: *The number of rating transitions that the yearly Unconditional Transition Matrix presented in Table 1 is based on.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**1561**	76	10	0	2	0	0	0	4
AA	28	**5054**	555	32	2	2	0	0	18
A	6	248	**13574**	864	37	10	8	4	98
BBB	2	14	519	**11464**	512	82	12	6	165
BB	1	4	7	494	**5725**	493	36	10	100
B	0	0	1	22	456	**4361**	160	38	90
$CCC-C$	2	0	0	0	12	73	**196**	25	38

Table 7: *The number of rating transitions that the H matrix presented in Table 2 is based on.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**811**	46	8	0	2	0	0	0	2
AA	24	**2622**	197	22	2	2	0	0	16
A	2	130	**7268**	392	9	6	0	0	66
BBB	0	4	261	**6160**	248	24	0	0	79
BB	1	0	5	274	**3119**	251	6	2	80
B	0	0	1	12	256	**2385**	58	8	42
$CCC-C$	0	0	0	0	6	39	**124**	9	22

Table 8: *The number of rating transitions that the L matrix presented in Table 2 is based on.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**750**	30	2	0	0	0	0	0	2
AA	4	**2432**	358	10	0	0	0	0	2
A	4	118	**6306**	472	28	4	8	4	32
BBB	2	10	258	**5304**	264	58	12	6	86
BB	0	4	2	220	**2606**	242	30	8	20
B	0	0	0	10	200	**1976**	102	30	48
$CCC-C$	2	0	0	0	6	34	**72**	16	16

Table 9: *The standard deviations (in %) of the probabilities in the yearly Unconditional Transition Matrix presented in Table 1.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**0.56**	0.52	0.19	0.00	0.00	0.00	0.00	0.00	0.12
AA	0.09	**0.42**	0.39	0.10	0.02	0.02	0.00	0.00	0.07
A	0.02	0.11	**0.23**	0.19	0.04	0.02	0.02	0.01	0.07
BBB	0.01	0.03	0.17	**0.27**	0.17	0.07	0.03	0.02	0.10
BB	0.01	0.03	0.04	0.31	**0.45**	0.31	0.09	0.05	0.14
B	0.00	0.00	0.02	0.09	0.40	**0.50**	0.24	0.12	0.18
$CCC-C$	0.41	0.00	0.00	0.00	0.98	2.19	**2.66**	1.39	1.68

Table 10: *The standard deviations (in %) of the probabilities in the yearly H Matrix presented in Table 2.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**0.85**	0.76	0.32	0.00	0.16	0.00	0.00	0.00	0.16
AA	0.17	**0.54**	0.47	0.16	0.05	0.05	0.00	0.00	0.14
A	0.02	0.14	**0.30**	0.25	0.04	0.03	0.00	0.00	0.10
BBB	0.00	0.03	0.23	**0.35**	0.23	0.07	0.00	0.00	0.13
BB	0.03	0.00	0.06	0.43	**0.61**	0.41	0.07	0.04	0.24
B	0.00	0.00	0.04	0.13	0.55	**0.65**	0.27	0.10	0.23
$CCC-C$	0.00	0.00	0.00	0.00	1.21	2.80	**3.43**	1.47	2.21

Table 11: *The standard deviations (in %) of the probabilities in the yearly L Matrix presented in Table 3.*

	AAA	AA	A	BBB	BB	B	$CCC-C$	D	NR
AAA	**0.73**	0.69	0.18	0.00	0.00	0.00	0.00	0.00	0.18
AA	0.07	**0.64**	0.63	0.11	0.00	0.00	0.00	0.00	0.05
A	0.03	0.15	**0.35**	0.30	0.08	0.03	0.04	0.03	0.08
BBB	0.02	0.05	0.26	**0.41**	0.26	0.13	0.06	0.04	0.15
BB	0.00	0.06	0.05	0.46	**0.67**	0.48	0.17	0.09	0.14
B	0.00	0.00	0.00	0.13	0.57	**0.76**	0.42	0.23	0.29
$CCC-C$	0.96	0.00	0.00	0.00	1.64	3.50	**4.14**	2.59	2.59

B Results for perturbed thresholds

This appendix shows the results for slightly perturbed threshold values compared to the thresholds in Table 5.

Table 12: *Given the specified thresholds, the table shows how many times (in percent, out of 19) the L2 distance between the H or L matrix and the M matrix is smaller than the L2 distance between the UTM and the M matrix.*

	SPREAD	PMI	GDP
Threshold	1.1 (%)	49	2.5 (%)
Entire matrix (%)	63 (11.1)	63 (11.1)	63 (11.1)
Default column (%)	68* (10.7)	58 (11.3)	63 (11.1)
Entire matrix, weighted (%)	63 (11.1)	79* (9.4)	68* (10.7)
Default column, weighted (%)	74* (10.1)	53 (11.5)	68* (10.7)

The results in Tables 12 and 13 still show that using H and L matrices to approximate the transition matrix one period ahead is better than just using the aggregated unconditional transition matrix, UTM.

Table 13: *Given the specified thresholds, the table shows how many times (in percent, out of 19) the L2 distance between the H or L matrix and the M matrix is smaller than the L2 distance between the UTM and the M matrix.*

	SPREAD	PMI	GDP
Threshold	1.3 (%)	51	3.5 (%)
Entire matrix (%)	63	58	58
	(11.1)	(11.3)	(11.1)
Default column (%)	68*	53	58
	(10.7)	(11.5)	(11.3)
Entire matrix, weighted (%)	63	74*	74*
	(11.1)	(10.1)	(10.1)
Default column, weighted (%)	74*	47	63
	(10.1)	(11.5)	(11.1)