RATING MIGRATIONS: THE EFFECT OF HISTORY AND TIME

Huong Dang and Graham Partington

University of Sydney

Abstract

We use the Cox proportional hazard model to investigate the probability of rating transitions using data for the period 1986 to 2005. Variables that capture rating history and the current rating significantly affect the probability of a rating transition. Different models are required for upgrades and downgrades, but the evidence consistently shows a tendency for history to repeat itself. Longer lagged durations in ratings tend to lead to longer subsequent durations and rating changes exhibit momentum. In addition to lagged duration and the direction of the lagged rating change, other significant variables are the rate of prior rating changes, the firm’s first ever rating, the time elapsed since that first rating, and having a period of being unrated. There is also evidence of interactions between the time spent in a rating grade and the main effect variables. The extent of these time interactions is greater for downgrades than for upgrades. The nature of the interaction is that the impact of the rating history variables diminishes as the time spent in the current rating gets bigger. The time interaction for the current rating diminishes the impact of the current rating for downgrades and intensifies it for upgrades.

JEL classification: C13, C14, C32, C34, C41, G14

Keywords: Event study, survival analysis, proportional hazard, time-varying covariates, rating migration, rating history, non-Markovian behaviors

1 Corresponding author: Huong Dang. Email: d.huong@econ.usyd.edu.au. Tel.: +61-421523259; Fax: +61-2-93516461. The authors wish to thank Paul Allison for advice and the Capital Markets Cooperative Research Centre CMRC for financial support
1. Introduction

Credit ratings are widely used to assess the risk that a firm will default and the probability of rating changes is used in pricing debt and in risk management. The use of credit ratings for credit risk measurement and management is particularly important under the new capital adequacy framework of the BASEL II Accord. Consequently, the Accord has stimulated much interest in the modelling of rating migration for both risk management and capital adequacy purposes.

Modeling rating migration is facilitated by an understanding of the underlying rating migration dynamics. It has been common to assume that rating migration follows a Markov process. This is a convenient assumption as the migration probability will only depend on the current rating and the ending rating, while the history of the rating changes is irrelevant. However, there is empirical evidence such as Altman (1998) and Hamilton and Cantor (2004) that rating migrations do depend on rating history. One objective of this paper, therefore, is to estimate models for the probability of rating migration based on rating history. The contribution is three-fold, first in extending the evidence of non-Markovian behavior in rating dynamics, second in extending the variables from rating history that are candidates for predictive models, and third in applying the Cox proportional hazards model which has had little use in this context.\(^2\) We also use a more extensive range of rating grades than has been the practice in prior work.

A second objective of the paper is to examine how the effect of rating history interacts with time. While the rating continues in its current state the distance in time from the historic observations is extending. The impacts of history variables on the migration

---

\(^2\) The only study that we are aware of that uses the same technique in relation to rating migrations is Figlewski et.al. (2006).
probability are thus likely to become increasingly “stale”. Our point here is not just that
more distant variables are likely to be less relevant, but that the impact of the variables
interacts with the duration of the rating. We hypothesize that the variables’ impact
decays over the time for which the current rating persists. We are not aware of any other
work which focuses on this interaction between rating history variables and time.

The study addresses the foregoing issues by examining the probable duration of rating
grades. How probable is it that by time $t$ (in the set $t = 1$ to $n$) there will be an upgrade
or a downgrade. It is possible to obtain this probability from a survival function $S(t)$
which can in turn be obtained from a hazard function estimated using Cox’s (1972)
proportional hazards model. Where $S(t) = P(T > t)$ and $P(T > t)$ is the probability that the
time of the rating transition $T$ will be after time $t$.

The results show that the several rating history variables are significant predictors of
future rating transitions and that there are some aspects of history repeating itself.
Furthermore, for both upgrades and downgrades, the impact of influential rating
transition variables is time-dependent. The impact of rating history decays the longer a
rating continues unchanged. The impact of the current rating is also time dependent, for
downgrades the effect diminishes with rating duration, but for upgrades it intensifies.
The time interactions are more strongly evident for downgrades relative to upgrades.

The paper is structured as follows: Section 2 provides an overview of the literature and
discusses the research questions; Section 3 describes the research method; Section 4
identifies the data; Section 5 presents the results of the analysis; Section 6 discusses the
robustness checks on the results and Section 7 summarizes the main findings.

2. Literature Review and Research Questions

2.1. Literature review
Previous empirical studies have found evidence of non-Markovian behaviour in rating migrations, such as serial correlation, momentum, and duration dependence. Evidence of serial correlation in rating migrations was provided by Altman and Kao (1992) and Carty and Fons (1994). Hamilton and Cantor (2004) also show that the direction of a prior rating change impacts on the current migration probability. Figlewski et al (2006) provide evidence of rating momentum, a downgrade is more likely to be followed by a downgrade than an upgrade.

The existence of a negative relation between the migration probability and the length of time a bond stays in a particular rating was suggested by Lando and Skodeberg (2002). This evidence of duration dependence helps motivate our use of the Cox proportional hazards model, since an attractive feature of the model is that it controls for underlying duration dependence, without requiring specification of the functional form of that dependence.

Newly rated firms, compared with seasoned firms of the same rating class, have a smaller probability of rating migrations within a few years, Altman (1998). Figlewski et al (2006) show that the longer the time elapsed since a firm was first rated the more likely it is that the firm will default.

Studies such as Holthausen and Leftwich (1986) have shown that there is a differential stock price response to upgrades and downgrades, with downgrades having a bigger effect than upgrades. A possible explanation for this is that downgrades may be more difficult to forecast than upgrades and may follow a different stochastic process. This is one factor that motivated the decision to develop separate models for upgrades and downgrades.

2.2 Research questions and hypotheses
Taking a Markovian approach to rating history, all of the relevant information about the impact of history is captured in the firm’s current rating, which we call the start rating since it defines the beginning of the current rating state. One question is whether the start rating alone is significant in explaining rating migrations, or whether additional rating history variables are significant. If the additional history variables are significant another question is whether there is an interaction between these variables and time. We hypothesize that the impact of rating history variables decay as the current rating continues on through time. If so, interaction variables for history with time should reduce the impact of the main effect variables for history.

Based in part on the research discussed in the literature review, we hypothesize that there is a tendency for rating history to repeat itself and that recent history has the strongest effect. If so the direction of the directly prior rating re-grades will positively affect the probability of a further re-grade in the same direction and negatively affect the probability of a re-grade in the reverse direction. Additionally, the longer the duration of the directly prior rating states the longer the likely duration of the current rating state. Any duration effect may, however, be conditional on the re-grades being in the same direction. The disruption in continuity created by re-grades in opposite directions could disrupt persistence in duration.

The effect of recent history on duration is expected to diminish with the passage of time and if so the penultimate (lag-one) duration will have a stronger impact on the migration hazard than the antepenultimate (lag-two) duration. We also hypothesize that a history of frequent rating changes is likely to be repeated. Thus, the higher the rate of prior rating migrations the higher the hazard of future rating migrations.

A natural question to ask is from how far back is the impact of rating history felt? The answer according to Figlewski et al (2006) is that it extends right back to the firm’s first
rating. Thus the original rating is incorporated in our model and the question is whether its effect remains in the presence of the additional rating history variables that our model also includes. According to Altman (1998) and Figlewski et al (2006), the age since first rating is also expected to impact on rating migrations and the effect is expected to be negative.

A period of being unrated creates a break in the rating history of the firm. However, it is not clear what impact this may have on the probability of rating migrations once the firm becomes rated again. Some firms may withdraw from being rated because they no longer carry significant debt, but firms are likely to withdraw from being rated when they have poor ratings and/or expect downgrades. Thus firms are likely to become unrated when they are of poor credit quality. Such firms may subsequently restructure their business to improve their financial status. These firms may then decide to be re-rated when they are likely to receive good credit ratings. Alternatively, firms may seek re-rating when they need to make a debt issue, even if their rating has not improved. It is not clear therefore what a period of being unrated signals about current credit quality, or the probability of rating re-grades. However, if firms use becoming unrated as a repeat strategy to avoid downgrades, then being down-graded will have a reduced probability for such firms.

3. Method

3.1. Rating states

A rating state starts from the time the firm enters a rating class (start rating) subsequent to the commencement date of the study. The state ends at the time the firm either migrates to another rating class (ending rating), becomes unrated or the study ends. The time a firm keeps the same rating is the survival time. However, if a firm exits from a
rating class due to a merger, extinction of rated debt, or any reason other than an upgrade or a downgrade, the survival time is treated as censored. Rating runs commencing before the start of the model estimation period (1 January 1986) or finishing after the end of the model estimation period (31 December, 2000) are also treated as censored.³

The duration of each rating state, until transition or censoring, was measured. The completed transitions were then labelled as upgrades, downgrades, or censored, according to their ending rating states. These rating transitions were then pooled and as a result a firm may contribute several rating transitions to the data-set. The use of multiple observations for the same firm is likely to introduce dependence among the observations. However, this problem is diminished to the extent that the covariates in the model control for dependence. To get robust standard errors in the presence of any residual dependence we use the Wei-Lin-Weissfeld method (Wei, Lin, Weissfeld, 1989). This method, however, does not correct for any remaining bias in the coefficients.

Two models were developed, one for the probability of upgrades, the other for downgrades.⁴ The estimation procedure makes use of risk sets, which are composed of all the firm ratings that are at risk of a rating change at time \( t \). In the process of estimating the model a new risk set is formed at each event time \( t \) when a rating transition occurs. Firm ratings leave the risk set once they experience a rating transition, or when they are censored. In forming the risk sets for the upgrade model, downgrades are treated as censored and vice versa.

³ The Cox model uses both completed transitions and censored observations in the estimation process.
⁴ An all-run model would make no sense since different models with different signs on common variables result for upgrade runs and downgrade runs.
The first set of upgrade and down grade models were estimated without allowing for any interaction of the variables with time. These models were then extended by adding the time interaction terms. Robustness tests were conducted by estimating the models over different time periods and using a different random sample.

3.2. Estimation and variables

The Cox proportional hazards model works by estimating the hazard rate, which is the rate of change of the survival probability over an interval, conditional on survival until the start of that interval. The survival probability can then be derived from the hazard. The hazard model to be estimated is:

\[ h(Z,t) = h(0,t) \exp^{Z\beta} \] (1)

Where \( h(Z,t) \) is the hazard for a rating transition at time \( t \) given the covariate vector \( Z \).

\( h(0,t) \) is the baseline hazard, which is that hazard with the covariate vector set to zero.

\( \beta \) is the vector of estimated coefficients

The vector of covariates \( Z \) contains:

**Lag one**: The duration (in years) of the non-censored rating immediately preceding the current rating.

**Lag two**: The duration (in years) of the non-censored rating immediately preceding the lag one rating.

**Rate prior change**: This equals the number of rating changes observed between the entry of the firm to the study and the beginning of the current state divided by the period over which the changes were observed.
Original rating: the rating of the firm when it was first rated.

Start rating: The rating at the beginning of each rating state.

Age since first rated: The rating age of the firm, which is equal to the length in years from the time the firm was first rated until the beginning of the current state.

Dummy NR: This variable takes the value of one if the firm became not rated (NR) at any point from the time it entered the study until the beginning of the current rating state, otherwise it is zero.

Dummy lag down: This variable captures the direction of the lag-one re-grade and takes the value of one if the lag one rating ends with a downgrade and zero otherwise.

Industry Dummies: Firm’s industry sectors, as identified by Standard & Poors, were used as control variables. The industry dummy took a value of one if the firm was in an industry sector and zero otherwise. Firms in the financial institution sector were excluded from the study, which left twelve sectors in the study. The twelve industry sectors are given in Table 1 and resulted in eleven dummy variables with the insurance sector left un-coded.

Time interactions: These terms are created through multiplication of each rating history variable by the event time $t$, where the event time $t$ is updated as each risk set is formed. Consequently, the interaction terms are time varying and the covariate vector becomes, $Z_t$. As rating age (Age since first rated) and rating volatility (Rate prior change) are functions of time, we do not examine the interactions between these variables and time. The interaction terms are:

$\text{Lag one time} = \text{Lag one} \times \text{Event time}$
Lag two time = Lag two * Event time

Start rating time = Start rating*Event time

Original rating time = Original Rating* Event time

NR time = Dummy NR * Event time

Lag down time = Dummy lag down * Event time

Because some of the covariates are time varying, the proportionality property of the Cox model no longer holds. This poses no problem for the estimation of coefficients on the covariates in the model, but estimation of the baseline hazard becomes problematic. Without the baseline hazard it is not possible to form the survival function and estimate the survival probabilities. For this reason we do not make a hold-out sample assessment of the predictive accuracy of the models.⁵

Unlike most studies on rating dynamics, which focus on coarser rating categories, such as AAA, AA, or investment and speculative grades, we employ the full rating sub-categories such as AAA, AAA-, AA+, AA, AA-. The rating scales are coded from 0 to 26 with 0 indicating the default state (D) and 26 indicating the AAA state. Details of the rating codes are provided in Table 1. A similar coding technique was employed by Kim, and Wu (2006) to examine the impact of sovereign credit ratings history on international capital inflows to emerging economies, and on the development of the financial sectors in these economies. The numeric conversion maintains the rank order of the rating and assumes that the difference between any two consecutive rating states is the same. For instance, it is assumed that the “rating gap” between BB+ (15) and BBB- (16) is the same as “rating gap” between C- (1) and C (2). While this might not be the case, the alternative of coding each rating class through dummy variables would consume a substantial

⁵ Our ongoing research is directed to solving this problem.
number of degrees of freedom. Adding an extra twenty-five dummy variables to the model would also preclude compact presentation of the results and make interpretation rather difficult, particularly in the case of the models where we introduce interaction terms.

TABLE 1 HERE

4. Data

Rating data was obtained from Standard & Poors CreditPro 2005. The whole dataset includes the rating history of 11,605 firms (of which 63.2% are American) over the period 1981-2005. A random sample of 3000 firms was selected from the period 1 January 1986 to 31 December 2000 and this was used to estimate the base (generic) model. In subsequent robustness tests we subdivided the estimation period into two parts and take a further random sample, we also estimate the model on the data for 2001 to 2005.

The year 1986 was used as the starting point of our study as the high yield bond market in the US was being established in the first half of the 1980’s. Rating migrations are more common events in the high yield bond sector. Thus migrations post 1985, consequent to the establishment of the high yield bond market, were expected to constitute a source of events for the study.

This time span of 15 years is long enough to cover different phases of the business cycle, major market downturns and international crises. The estimation period witnessed the US stock market crash in 1987, the Mexican currency crisis in 1994, the Thai financial crisis in 1997, the Russian sovereign bond default in 1998, and the collapse of Long Term Capital Market (LTCM) hedge fund in 1998.

Histograms of rating durations are depicted in Figure 1. Both upgrades and down grades have positively skewed distributions. The range of the distributions is similar, but it is
clear that durations for downgrades tend to be shorter than for upgrades. There is a noticeable concentration of downgrades in durations shorter than one year.

FIGURE 1 HERE

Table 2 provides descriptive statistics for the samples of rating durations across the full period and in the sub-periods used for the robustness checks. The full sample used in the upgrade model consists of 1,113 ratings (37.1%) experience migrations (downgrades), and 1,887 ratings (62.9%) are censored (including upgrades). Downgrade ratings vary from 1 day to 11.17 years, have a mean length of 1.49 years, and a median length of 0.909 years.

For the full sample used for the upgrade model, 726 runs (24.2%) experience migrations (upgrades) and 2274 runs (75.8%) are censored (including downgrades). Upgrade runs vary from 2 days to 10.44 years, have mean length of 2.19 years, and a median length of 1.75 years.

TABLE 2 HERE

5. Results

The model given by equation 1 was estimated for the upgrades and for the downgrades. The results of the base (generic) models and the models extended by the time interaction variables are given for down-grades in Panel A - Table 3 and for up-grades in Panel A - Table 4. Panels B - Tables 3 and 4 provide statistics on the fit of the respective models.

In interpreting Table 3 and Table 4, a negative coefficient reduces the hazard and therefore reduces the probability of a rating migration. The reported hazard ratios represent the relative change in the hazard for a one unit change in the independent variable. For example, in the base model for downgrades, covering 1986-2000 (Panel A
– Table 3), an increase in the length of the lag one rating run by one year reduces the chance of a downgrade by \((1 - 0.937)\) or 6.3%.

The results in Tables 3 and Table 4 are generally as hypothesized. The hazard of a rating change depends significantly upon several aspects of rating history as well as the current rating. There is some tendency for history to repeat itself. For example, longer lagged durations increase the probability of the current rating continuing as do lagged re-grades in the same direction, and more frequent re-grades tend to make a current downgrade more likely. It is also clear that there are significant interactions with time for most rating history variables, and the current (start) rating. The interaction effects are generally consistent with decay in the impact of rating history as the longer a rating remains unchanged. The impact of the current rating also decays with rating duration for downgrades, but intensifies with rating duration for upgrades. While the foregoing provides a general picture of the results, there are some differences between the upgrade and downgrade models.

5.1. Results for 1986-2000 for downgrades

The results for the full estimation sample for the downgrade model (the first six columns of Panel A - Table 3), show that, a longer lagged duration at lag one (but not lag two), a higher original rating, and a higher start rating for the current observation all significantly reduce the probability of a rating downgrade, although the original rating is only marginally significant. It is interesting to observe that a prior break in rating history also reduces the probability of a downgrade. The probability of a downgrade is significantly increased by a higher rate of prior rating changes and a downgrade at lag one. In contrast to the results of Altman (1998) and Figlewski et al (2006) the rating age of the firm has no significant effect. The impact of a downgrade at lag one is particularly strong. As shown by the hazard ratio (column three of Panel A - Table 3)
relative to an upgrade at lag one, a downgrade at lag one, results in two and a half times the risk of a further downgrade.

Comparison of the log-likelihood statistics in Panel B - Table 3 shows that adding the time interaction terms improves the model. A likelihood ratio test (not reported) shows that this improvement is significant at better than the one percent level. With the exception of the original rating, which becomes insignificant, adding the time interaction terms does not change the significance of the main effects. However, the absolute values of the main effect coefficients increase. The increase is modest in most cases, but it is substantial in the case of the dummy for a downgrade at lag one. As a consequence, the hazard ratio, given a downgrade at lag one, rises above four. Note however, that this large effect only applies at the start of the rating. As time passes the time interaction kicks in and the impact of the lagged re-grade progressively reduces.

For each of the main effects that is significant, the corresponding interaction term is also significant. The significant interaction terms all have coefficients that are of the opposite sign to the main effect. Thus, for downgrades the impact of the rating history and the current rating grows less the longer the rating continues unchanged.

TABLE 3 HERE

5.2. Results for 1986-2000 for upgrades

The results of the upgrade model for the full estimation sample (the first six columns of Panel A - Table 4) show that longer durations at lags one and two, a higher rating for the current run, a break in rating history by being unrated, and a downgrade at lag one all significantly reduce the probability of an upgrade. In contrast, a longer period from first being rated significantly increases the probability of an upgrade. This latter result is the reverse of the results of Altman (1998) and Figlewski et al (2006).
In contrast to the downgrading model, the rate of prior rating changes and the original rating have become insignificant, while the duration at lag two and the period since the first rating have become significant. The dummy for a downgrading at lag one has changed sign, but this is to be expected. The effect of this variable is particularly strong, as was the case for the downgrading model. The hazard of a current upgrade, given a downgrade at lag one, is sixty-two percent of the hazard of cases where there was an upgrade at lag one. In other words, with an upgrade at lag one a further upgrade is $1/0.62 = 1.6$ times more likely.

The improvement in the likelihood statistics (Panel B - Table 4) from adding the time interactions is noticeably less than was for the case for downgrades. Nevertheless, the improvement is significant at better than the five percent level. With the exception of the duration at lag two, which becomes insignificant, adding the time interaction does not change the significance of the main effects. Neither is there much change in the estimated coefficients, except in the case of the dummy for a downgrading at lag one. Here the coefficient change is such that an upgrade at lag two makes a further upgrade more than twice as likely before the interaction with time starts to diminish the effect.

Only two of the time interaction terms are significant. One of these is the interaction with the dummy for a downgrading at lag one, and the coefficient has the opposite sign from the main effect. Thus the impact of the re-grades at lag one decays with time. The other significant interaction term is with the current rating. In this case the interaction coefficient has the same sign as the main effects, thus intensifying the main effect as the current duration extends. Consequently, there is a further reduction in the hazard of an upgrade the longer the current rating continues.

6. Robustness check
6.1. The estimation periods

The robustness tests were designed to address the question do the effects of rating history behaviors and time interactions remain significant under different market conditions? The robustness tests were conducted over three time periods, 1986 to 1990 and 1991 to 2000 and 2001 to 2005. Each period saw market dramatic changes in capital markets, but they were of a markedly different nature.

The period 1986-1990 witnessed the deregulation in the US savings and loan industry, coupled with changes in the US tax policies. This period also marked the beginning of the risk based capital regulation framework. The US stock markets crashed in 1987 and most other stock markets also crashed. Overall, this period saw short term turbulence in the equity markets and ended with the housing bubble in the UK.

The period 1991-2000 witnessed historically low, stable (or declining) interest rates and inflation in the US and other developed markets. The period started with a recessionary year in the US followed by a long expansion and a bull market until 1999. In contrast to the expansion in the US, several international crises occurred with profound effects spreading globally. These include the Mexican peso collapse of 1994, the Thai financial crisis of 1997, the Russian sovereign bond default in the summer of 1998, and the collapse of the LTCM hedge fund in the same year.

The 2001-2005 period saw the Internet bubble burst and the 9/11 terrorist attack. The bursting of the Internet bubble was followed by an economic slowdown and falling business investment in the US, with negative stock returns for three consecutive years 2000 to 2002. The year 2002 was considered the worst year for the corporate bond market in over 20 years. This year saw unprecedented credit deterioration and the dramatic bankruptcies of fallen angels like WorldCom. Of the ten biggest bankruptcies

Corporate rating volatility intensified and the default rate escalated.

The three samples for robustness tests were constructed as follows:

The 1 January 1986 to 31 December 1990 dataset includes 911 ratings, of which 24.9% experience downgrades and 16.2% experience upgrades. Due to the smaller sample size relative to other periods, the whole dataset of 911 ratings was used.

For the 1 January 1991 to 31 December 2000 period, a random sample of 3000 ratings was taken. Within this sample, 33.1% of ratings experience downgrades and 19.5% experience upgrades.

For the 1 January 2001 to 31 December 2005 period, a random sample of 3000 ratings was taken. Downgrades represent 38.8% of the sample and upgrades are 14.7%.

In comparing the incidence of re-grades across the periods it should be borne in mind that shorter periods will tend to have more censored observations and this will tend to depress the observed incidence of re-grades. In the discussion of upgrade and downgrade models that follows, the comparison of coefficients is undertaken across the three periods used for the robustness test and the original estimation sample.

6.2. Downgrade models

Panel A - Table 3 shows that rating history variables, and the current rating, affect the probability of a downgrade in each period. It is also evident that time interactions reduce the impact of the main effects in each period. However, there is some variation in the significant variables and the magnitude of coefficients across the different periods.

With the exception of the duration at lag two, all the significant variables have the same sign across all periods. However, the duration at lag two, the rate of prior rating change,
the original rating, and the period since first rated are significant in some cases, but not others. Larger values for these variables generally reduce the hazard of a rating downgrade.

In all cases the duration at lag one and a break in rating history significantly reduce the hazard of a downgrade. In contrast, a downgrade at lag one significantly increases the risk of a subsequent downgrade in all cases. This variable consistently has the strongest impact, reaching a maximum hazard ratio of over ten when time interactions are included in the model for the 1986 to 1990 period.

With the introduction of the time interaction variables the improvement in the log-likelihood statistics is significant over all periods (Panel B – Table 3). All of the interaction terms are significant, but not in all periods. The interactions for the duration at lag two, the original rating, and a period of being unrated, are significant in some periods, but not others. The duration at lag one, the current rating, and the direction of the re-grade at lag one are significant in every period. The sign of the significant coefficients for the interactions is, in all cases, the opposite of the sign for the main effect, consistent with a decay in impact over time.

6.3. Upgrade models

Panel A - Table 4 shows that rating history variables, and the current rating, affect the probability of an upgrade in each period. It is also evident that time interactions reduce the impact of the main effects of rating history in each period. Relative to the downgrade model, however, there is less consistency in the results and there is less evidence of significant time interactions.

With the exception of the duration at lag two and the period since first rated the sign of the significant variables is the same across estimation periods. However, only one
variable, the start rating is significant in all cases, with a higher start rating reducing the probability of an upgrade. A downgrade at lag one and a period unrated have the strongest effects in reducing the probability of an upgrade, but a downgrade at lag one is not significant in the 1986 to 1990 period and being unrated is not significant in this period after controlling for time interactions.

After introducing the time interaction variables there is a significant improvement in the likelihood ratio, but the improvement is less significant than for the downgrade model (Panel B – Table 4). Only a few of the interaction terms are significant and in the case of the rating history variables they have the opposite sign to the main effects.

7. Conclusion

A Cox regression model is used to estimate dynamic models for the hazard of rating migrations. The purpose of this modeling is to investigate whether a set of rating history variables, in addition to the current rating, are significant determinants of the probability of a rating migration. That is to say we test for non-Markovian behavior in rating migrations. The study also investigates whether the effect of the predictor variables is constant, or whether it interacts with the time elapsed in the current rating state.

Using a sample from Standard & Poor’s CreditPro 2005 dataset, hazard models were estimated for rating upgrades and downgrades for the period 1986-2000. Robustness checks of the models were conducted over two sub-periods (1986-1990 and 1991-2000) drawn from the estimation period, and a third period subsequent to the estimation period, 2001-2005. These robustness checks confirm the main findings discussed below. Apart from the rating duration at lag two and the period since first rated, the signs of the significant coefficients are unchanged across periods. However, there is variation across time in the significance and magnitude of some of the estimated coefficients. Changes
in significance are particularly evident for the original rating, the rating duration at lag two and the time elapsed since the firm was first rated. The variation is sufficient to suggest that there is some dependence on the conditions prevailing in each period.

Overall, the results show that the hazard of a rating change depends on both the current rating and the history of the rating. The models differ between upgrades and downgrades, but they consistently show that history has a tendency to repeat itself. Longer lagged durations, tend to increase subsequent durations by reducing the probability of a current re-grade. If the lagged rating change was a downgrade, the probability of a current downgrade is increased and if the lagged rating change was an upgrade the probability of a current upgrade is increased. The impact of these lagged rating changes has a particularly large effect on the hazard of the subsequent rating change. Thus, there appears to be substantial momentum in rating changes.

The probability of either an upgrade, or downgrade, is reduced by having a high current rating, or by having experienced a period of being unrated. Having a high original rating reduces the probability of a downgrade and increases the probability of an upgrade, but this effect is only significant for some time periods. A higher level of past rating changes increases the chance of a downgrade in some periods, but does not significantly affect upgrades.

There is consistent evidence of an interaction between time and the main effect variables, although not for all variables in all periods. The extent of time interactions is greater for downgrades than for upgrades. The nature of the interaction is such that the impact of the rating history variables diminish as the time spent in the current run gets bigger. The impact of the current rating diminishes with time for downgrades and intensifies for upgrades.
References


Kavvathas, D. “Estimating credit rating probabilities for corporate bonds.”

Kim, S., E. Wu (2006). “Sovereign credit ratings, capital flows and financial sector development in emerging markets.”


Fig. 1: Histogram of state length in estimation sample
<table>
<thead>
<tr>
<th>Description</th>
<th>Codes/ Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date the firm was first rated</td>
<td>First rated date</td>
</tr>
<tr>
<td>The starting date / ending date of each rating state</td>
<td>Start_date / End_date</td>
</tr>
<tr>
<td>The length of a rating state</td>
<td>Duration</td>
</tr>
<tr>
<td>Rating age (since it was first rated ) at state entry</td>
<td>Age since first rated</td>
</tr>
<tr>
<td>The starting rating at the beginning of each rating state</td>
<td>Start rating</td>
</tr>
<tr>
<td>The original rating when the firm was first rated</td>
<td>Original rating</td>
</tr>
<tr>
<td>The length of the non-censored lag one rating state</td>
<td>Lag one</td>
</tr>
<tr>
<td>The length of the non-censored lag two rating state</td>
<td>Lag two</td>
</tr>
<tr>
<td>A measure of rating volatility</td>
<td>Rate prior change</td>
</tr>
<tr>
<td>Dummy variable indicating whether the firm underwent a Not Rated (NR) status during the time it spent in the study</td>
<td>Dummy NR</td>
</tr>
<tr>
<td>Dummy variable indicating whether the non-censored immediate prior rating state was a down state</td>
<td>Dummy lag down</td>
</tr>
<tr>
<td>Firm's sector coded as a dummy variable</td>
<td>Sector **</td>
</tr>
<tr>
<td>Insurance</td>
<td>Aerospace / automotive / capital goods / metal</td>
</tr>
<tr>
<td>Forest and building products / homebuilders</td>
<td>Consumer / service sector</td>
</tr>
<tr>
<td>Leisure time / media</td>
<td>Energy and natural resources</td>
</tr>
<tr>
<td>Health care / chemicals</td>
<td>Real Estate</td>
</tr>
<tr>
<td>Transportation</td>
<td>Telecommunications</td>
</tr>
<tr>
<td>Utility</td>
<td>High technology/ computers/ office equipment</td>
</tr>
</tbody>
</table>

* The beginning date of the study is 1 January, 1986. The estimation sample includes 3000 rating runs within the period 1 January, 1986 - 31 December, 2000

** 13 Sector categories were provided by Standard & Poor’s in CreditPro 2005 dataset. Financial institutions were excluded from the sample
<table>
<thead>
<tr>
<th>Sample</th>
<th>Type of run</th>
<th>Number of runs (percentage of sample size)</th>
<th>Number of runs with a prior NR</th>
<th>Number of runs with a lag one down run</th>
<th>Mean (years)</th>
<th>Median (years)</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Min (days)</th>
<th>Max (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986-2000</td>
<td>Down-run</td>
<td>1113 (37.1%)</td>
<td>131 (11.77%)</td>
<td>813 (73.05%)</td>
<td>1.496</td>
<td>0.9095</td>
<td>1.718</td>
<td>2.182</td>
<td>6.443</td>
<td>1</td>
<td>11.17</td>
</tr>
<tr>
<td></td>
<td>Up-run</td>
<td>726 (24.2%)</td>
<td>103 (14.19%)</td>
<td>284 (39.12%)</td>
<td>2.199</td>
<td>1.75</td>
<td>1.76</td>
<td>1.688</td>
<td>3.49</td>
<td>2</td>
<td>10.44</td>
</tr>
<tr>
<td>1986-1990*</td>
<td>Down-run</td>
<td>227 (24.91%)</td>
<td>14 (6.17%)</td>
<td>181 (79.73%)</td>
<td>0.895</td>
<td>0.63</td>
<td>0.798</td>
<td>1.02</td>
<td>0.595</td>
<td>5</td>
<td>4.096</td>
</tr>
<tr>
<td></td>
<td>Up-run</td>
<td>148 (16.25%)</td>
<td>9 (6.08%)</td>
<td>70 (47.29%)</td>
<td>1.27</td>
<td>1.079</td>
<td>0.77</td>
<td>0.966</td>
<td>0.729</td>
<td>6</td>
<td>4.096</td>
</tr>
<tr>
<td>1991-2000</td>
<td>Down-run</td>
<td>992 (33.07%)</td>
<td>94 (9.47%)</td>
<td>767 (77.52%)</td>
<td>1.1055</td>
<td>0.619</td>
<td>1.216</td>
<td>1.607</td>
<td>2.82</td>
<td>1</td>
<td>6.88</td>
</tr>
<tr>
<td></td>
<td>Up-run</td>
<td>585 (19.5%)</td>
<td>63 (10.77%)</td>
<td>211 (36.07%)</td>
<td>1.817</td>
<td>1.471</td>
<td>1.319</td>
<td>1.207</td>
<td>1.23</td>
<td>6</td>
<td>7.35</td>
</tr>
<tr>
<td>2001-2005</td>
<td>Down-run</td>
<td>1163 (38.76%)</td>
<td>23 (1.97%)</td>
<td>1095 (94.15%)</td>
<td>0.464</td>
<td>0.246</td>
<td>0.577</td>
<td>2.657</td>
<td>9.329</td>
<td>1</td>
<td>4.337</td>
</tr>
<tr>
<td></td>
<td>Up-run</td>
<td>442 (14.73%)</td>
<td>11 (2.49%)</td>
<td>222 (50.22%)</td>
<td>1.189</td>
<td>1.051</td>
<td>0.839</td>
<td>0.869</td>
<td>0.523</td>
<td>1</td>
<td>4.69</td>
</tr>
</tbody>
</table>

* sample 1986-1990 includes 911 runs while each of the other samples includes 3000 runs

Table 2: Descriptive statistics of run length in samples
Panel A: Model summary across periods

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Generic Model</td>
<td>Extended Model</td>
<td>Generic Model</td>
<td>Extended Model</td>
<td>Generic Model</td>
</tr>
<tr>
<td></td>
<td>Parameter estimate</td>
<td>Standard error</td>
<td>Hazard ratio</td>
<td>Parameter estimate</td>
<td>Standard error</td>
</tr>
<tr>
<td>Lag one</td>
<td>0.04197*</td>
<td>0.03123</td>
<td>1.043</td>
<td>0.4837*</td>
<td>0.16614</td>
</tr>
<tr>
<td>Lag two</td>
<td>0.01106</td>
<td>0.01307</td>
<td>1.011</td>
<td>0.1457</td>
<td>0.1482</td>
</tr>
<tr>
<td>Original rating</td>
<td>0.00328</td>
<td>0.00695</td>
<td>0.997</td>
<td>0.057**</td>
<td>0.02702</td>
</tr>
<tr>
<td>Start rating</td>
<td>0.00247</td>
<td>0.00946</td>
<td>1.044</td>
<td>0.07456*</td>
<td>0.10968</td>
</tr>
<tr>
<td>NR time</td>
<td>0.17258**</td>
<td>0.05261</td>
<td>1.136</td>
<td>1.02019**</td>
<td>0.29671</td>
</tr>
<tr>
<td>Lag down_time</td>
<td>0.30274**</td>
<td>0.05407</td>
<td>0.739</td>
<td>-0.134**</td>
<td>0.37061</td>
</tr>
</tbody>
</table>

Panel B: Model fit statistics over different periods of time

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2Log L</td>
<td>DF</td>
<td>P value</td>
<td>-2Log L</td>
</tr>
<tr>
<td>Generic down-grade model</td>
<td>14484.19</td>
<td>&lt;.0001</td>
<td>2381.3</td>
<td>19</td>
</tr>
<tr>
<td>Extended down-grade model with time-varying variables</td>
<td>14319</td>
<td>&lt;.0001</td>
<td>2305.8</td>
<td>25</td>
</tr>
<tr>
<td>Difference between two models</td>
<td>165.14</td>
<td>6</td>
<td>&lt;.0001</td>
<td>75.42</td>
</tr>
</tbody>
</table>

*p ≤ 1% based on Wald chi-square tests

1%< p < 5% based on Wald Chi-square tests

**5%< p < 10% based on Wald Chi-square tests
Table 4: Regression models for up states across periods

### Panel A: Model summary across periods

#### Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Parameter estimate</th>
<th>Standard error</th>
<th>Hazard ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag one time</td>
<td>0.01715</td>
<td>0.01468</td>
<td>1.01</td>
</tr>
<tr>
<td>Lag two time</td>
<td>0.00453</td>
<td>0.02053</td>
<td>1.00</td>
</tr>
<tr>
<td>Original rating time</td>
<td>0.00293</td>
<td>0.00648</td>
<td>1.09</td>
</tr>
<tr>
<td>Start rating time</td>
<td>-0.01788</td>
<td>0.00808</td>
<td>0.92</td>
</tr>
<tr>
<td>NR time</td>
<td>0.00642</td>
<td>0.03733</td>
<td>0.91</td>
</tr>
<tr>
<td>Lag down time</td>
<td>0.1009**</td>
<td>0.04489</td>
<td>1.10</td>
</tr>
</tbody>
</table>

#### Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2Log L</td>
<td>DF</td>
<td>P value</td>
<td>-2Log L</td>
</tr>
<tr>
<td>Generic up-grade model</td>
<td>9334.5</td>
<td>19</td>
<td>&lt;0.001</td>
<td>1516.2</td>
</tr>
<tr>
<td>Up-grade extended model with time-varying variables</td>
<td>9318.25</td>
<td>19</td>
<td>&lt;0.001</td>
<td>1503.25</td>
</tr>
<tr>
<td>Difference between two models</td>
<td>16.456</td>
<td>0.0115</td>
<td>0.0049</td>
<td>13.067</td>
</tr>
</tbody>
</table>

* p ≤ 1% based on Wald chi-square tests
** 1%< p ≤ 5% based on Wald Chi-square tests
*** 5%< p ≤ 10% based on Wald Chi-square tests