MODELING RATING MIGRATIONS

Huong Dang and Graham Partington[1]

University of Sydney

30 July, 2007

Abstract

Using the framework of survival analysis and data from Standard & Poor's CreditPro 2005 dataset, we develop Cox proportional hazard models to estimate time-varying rating migration probabilities. We examine the duration of runs in rating grades, for example how long a firm stays in a rating grade before an upward revision (an up-grade run) or before a downward revision (a down-grade run). We find that there is statistically significant dependence of time-varying migration hazard on rating history prior to the rating run analyzed.

The cumulative accuracy profile (CAP) curve is used to assess the predictive accuracy of the probabilistic migration forecasts. The results show that the down-grade model and especially, the default model exhibit good accuracy in assigning the lowest survival estimates to the riskiest firms.

JEL classification: C13, C14, C32, C34, C41, C53, G12, G14

Keywords: Survival analysis, proportional hazard, rating migration, rating history, predictive accuracy, cumulative accuracy profile

[1] Corresponding author: Graham Partington. Email: g.partington@econ.usyd.edu.au, Tel: +61-2-90369429, Fax: +61-2-93516461.
The authors wish to thank Paul Allison and Vu Nguyen for advice and programming assistance respectively.

1. Introduction

Accurately modelling credit risks based on rating migration probabilities has become an important issue since the new Basel Accord Framework came into effect. The new Framework, with emphasis on rating migration probability, imposes strict requirements on bank's capital adequacy and the classification of securities. Under the new framework, banks are required to consider rating migration matrices published by rating agencies such as Moody's, Standard & Poor's in determining capital adequacy.

In light of the new Basel Accord Framework and current business practice, credit risk models have incorporated rating migration matrices to quantify downgrade risks and default risks. It is important to model rating migrations accurately as credit pricing can be sensitive to a small change in the estimated migration probability. Such modelling requires an understanding of rating migration dynamics. Furthermore, an understanding of rating dynamics is useful to predicting price movements in both bond and equity markets. Bond re-ratings can affect the price of equity securities as well as bonds, and price effects may spread to rivals (Caton and Goh (2003), Dichev and Piotroski (2001)).

Rating migration dynamics have been widely investigated in the finance literature during the past twenty years. There have been some findings of rating momentum and path dependence[2]. Given the observations of path-dependent behavior in rating migrations, it makes sense to raise the question whether and how rating history including lagged rating duration, rating age and rating volatility can explain time-varying rating migration probabilities. In this study Cox's proportional hazard model is used to address the following:

i. Estimate time-varying rating migration probabilities

ii. Investigate whether the time-varying migration hazard depends on:

- Rating history prior to the rating run analyzed.
- Countries in which firms are located , and
- Sectors in which firms operate

[2] See Atlman and Kao (1992), Carty and Fons (1994), Lando and Skodeberg (2002), Hamilton and Cantor (2004), Figlewski, Frydman and Liang (2006)

iii. Assess the predictive accuracy of probabilistic rating forecasts

Figlewski, Frydman, and Liang (2006) also employed the Cox proportional hazard models and used Moody's Default Research Database to examine the hazard of credit events such as default and major rating revisions over the period 1981-2002. They incorporated some rating history variables and a number of macroeconomics variables in their hazard models. There are several differences between our study and the study by Figlewski et. al.

Firstly, an important part of our study is to assess the predictive accuracy of the estimated model using a holdout sample over the period 2001-2005. The cumulative accuracy profile (CAP) curve is used to evaluate the accuracy of the estimated probability of a rating change. Figlewski et. al. (2006) did not conduct a test of predictive accuracy. Instead, they test the robustness of the estimated models over different sample periods.

Secondly, we focus on a number of non-Markovian behaviors in rating dynamics whereas Figlewski et. al. (2006) focus more on macroeconomic impacts on rating migrations. We also attempt to model whether the rating migration hazard (upgrade hazard, downgrade hazard, default hazard) depends on the country where a firm is located, and the sector in which a firm operates. Figlewski et. al. (2006) did not investigate country and sector impacts on rating migrations.

Thirdly, we consider every rating revision, which includes both minus and plus changes in rating grades, over the period 1986-2000. In contrast, Figlewski et. al. (2006) merged ratings into broad classes such as AAA, AA, A, BBB being combined into an investment class. Finally, there are differences between the covariates that we use and those Figlewski et. al. (2006) used.

The paper is structured the following way. The next section provides an overview of literature review and research questions. The subsequent section presents the method used and describes the data. The following section summarizes the results of the estimated models and presents the CAP curves. The final section summarizes the main findings of this research.

2. Literature Review and Research Questions

2.1. Literature review

Previous empirical studies found evidence of non-Markovian behaviours in rating dynamics. For example, a downgrade is more likely to be followed by a downgrade than an upgrade. Rating migrations exhibit duration dependence, serial correlation, and path dependence. For instance, Atlman and Kao (1992), Carty and Fons (1994) provided evidence of serial correlation in rating migrations. With respect to duration dependence, Lando and Skodeberg (2002) proposed that a negative relation exists between migration probability and the length of time a bond stayed in a particular rating. Hamilton and Cantor (2004) suggested that the direction of the prior rating change impacts on the migration probability. Figlewski et al (2006) also suggested that rating momentum exists.

Altman (1998) found that newly rated firms, compared with seasoned firms of the same rating class, have smaller probability of rating migrations within a few years. Figlewski et al (2006) also provided evidence of an aging effect. Specifically, the longer the length of time since a firm was first rated, the more likely that the firm would default.

A study by Caton and Goh (2003) indicated that the effect of a rating downgrade spreads to the re-rated firm's rivals. Furthermore, market reactions are not confined to bond performance but also impact on stock prices. Rating migrations generate an asymmetrical impact on stock returns, that is, downgrades have a more pronounced impact on stock returns than upgrades do. Holthausen and Leftwich (1986) suggested that no abnormal return was recorded following up-grade announcements. However, abnormally low returns were observed within three months following downgrade announcements. Dichev and Piotroski (2001) reported that firms experienced downgrade announcements within the past year witness negative abnormal stock returns of around 10 to 14%. Dichev et al (2001) further indicated that small firms with low credit quality would suffer more from stock underperformance following rating downgrades. A study by Choy, Gray and Ragunathan (2006) on stock returns of Australian firms following credit rating revisions also showed that "downgrades contain price relevant information".

The differential stock price response to bond upgrades and downgrades suggests that downgrades may be more difficult to forecast than upgrades, and possibly follow a different stochastic process. Consequently, separate models are developed for upgrades and downgrades.

2.2 Research questions

Despite the evidence of serial correlation and path dependence in rating migrations, it seems that this finding has not been widely incorporated into the forecasts of rating migration probabilities. An accurate forecast framework is more likely to be constructed if deviations from Markov behaviours in rating dynamics are built into the models.

In this study, we investigate a number of non-Markovian behaviors in rating dynamics after controlling for duration dependence. Specifically, we address the following questions:

i. Is there a positive relation between the rate of prior rating changes and rating migration probability? It is expected that the higher the rate of prior rating changes, the higher the hazard of future migrations.

ii. Is there a negative relation between the duration of the lag one rating run, which is the non-censored run immediately preceding the current run, and migration probability? It is expected that the longer the lagged rating duration, the lower the hazard of migration. Similarly, is there a negative relation between the duration of the lag two rating run, which is the non-censored run immediately preceding the lagged one rating run, and migration probability? It is expected that the lag one rating run would have a stronger negative impact on rating migration probability than the lag two rating run would.

iii. What is the impact of the direction of the immediately prior rating run on rating migration probability? It is expected, for example, that a lag one down run would increase the rating migration probability for down runs, and reduce the migration probability for up runs

iv. What are the impacts of past rating classes on rating migration probabilities? We examine the impacts of the original rating (rating of the firm when it was first rated), entry rating (rating of the firm when it entered the study) and start rating (rating of the firm at the beginning of each rating run analysed). It is expected that the original rating and the entry rating would not exert as strong an impact on migration hazards as the start rating since the latter rating is the most recent. The start rating is expected to have a negative effect on the migration hazard, that is, the better the start rating, the lower the migration hazard. We are uncertain as to the impacts of entry rating and original rating on migration hazards for each type of run. Figlewski et al (2006) found the different impacts of the original rating on migration hazards for "fallen angels", "rising stars", and average firms.

v. What is the impact of a prior Not rated (NR) status on future rating migration hazard? This arises when the firm was not rated for some period of the study prior to the run currently analysed.

vi. What is the impact of age on rating migration probabilities? It is expected that the rating age would have a positive impact on migration hazard, that is, the older the firm since it was first rated, the higher the hazard of migrations. Altman (1998) and Figlewski et al (2006) found evidence of such an aging effect.

vii. What are the impacts of the country and the industry sector on rating migration hazard? These variables might have an influence on rating changes in their own right or they might proxy for other variables. For example, a highly levered industry might have a higher hazard of rating migration.

3. Method

3.1. Cox proportional hazard model

We adopt the framework of event history analysis (Allison, 1984) and the Cox proportional hazard model (Cox, 1972) to estimates the time-varying probabilities of rating migrations, and to investigate the effect of individual covariates on these probabilities. The Cox proportional hazard model (Cox, 1972) has been very popular in survival analysis in medicine, and it has been increasingly applied in finance studies.

The Cox proportional hazard model was used to estimate time-varying probabilities of stock price reversal (Yao, Partington, and Stevenson, 2005), and to study price reversal in the UK property market (Partington and Stevenson, 2001). In this study, we build upon the approach presented in these papers. Figlewski, Frydman, Liang (2006) also used the Cox proportional hazard model to examine the effects of firm-specific and macroeconomic factors on default and major rating migrations over the period 1981-2002. While their research emphasized the impacts of macroeconomic variables on defaults and major credit events, our study focuses on non-Markovian behaviors in a finer partition of rating migrations during 1986-2000. We also test the predictive accuracy of probabilistic migration forecasts over the period 2001-2005.

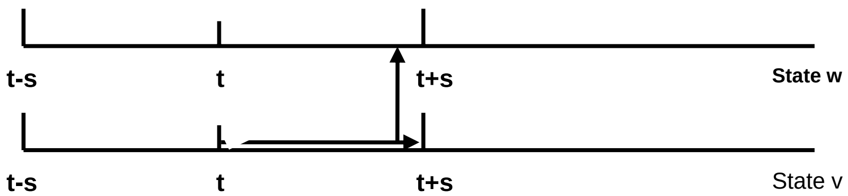
3.2. Rating runs and estimation

A rating run starts either from the time the firm enters a rating class (starting rating), or from the time the firm enters the study (whichever is later) until the time the firm either migrates to another rating class (ending rating), becomes unrated (NR) or comes to the end of the study. The time a firm keeps the same rating is considered the survival time.

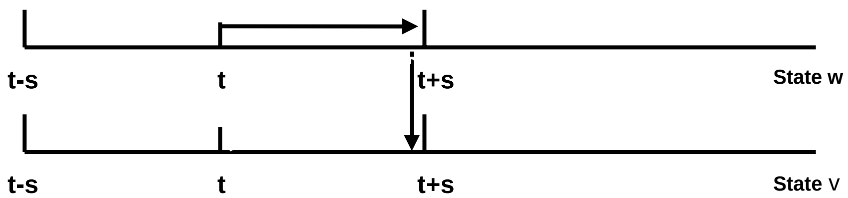
The survival function estimates the probability that the firm will keep its current rating class. For every survival function, there is a hazard function r(t) which measures the rate of change in the survival probability. The hazard function can be interpreted as an estimate of the number of migrations in a period. For instance, in this study a migration hazard of 0.1 indicates that there would be a tenth of a migration in a year or one migration in ten years. In the model the hazard is allowed to change through time as the rating run continues.

Assume s is the time a firm spends in its current rating v, and (t+s) is the time the firm migrates from rating v to rating w. This rating migration forms an event. If a firm exits from a rating class due to merger, extinction of firm's rated debt, or any reason other than an up-grade or a down-grade, 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.

A firm is considered to be upgraded when it migrates from rating v to rating w (better credit quality) at time (t+s). This constitutes an up-grade event at time (t+s).



Conversely, a firm is considered to be downgraded when it migrates from rating w to rating v (lower credit quality) at time (t+s). This constitutes a down-grade event.



An event occurs when a rating run ends due to either an up-grade or a down-grade. The rating runs for each firm in the dataset were treated as separate observations. Depending on the ending states, the rating runs were labeled as up-grade, down-grade and censored. These runs were then pooled for all firms, and hazard models were developed to estimate the migration probability.

The use of repeated events for the same firm is likely to introduce dependence among the observations. The dependence among multiple observations is due to unobserved heterogeneity in the baseline hazard rate (Allison, 1995). This problem is reduced to the extent that covariates in the model control for dependence and hence, reduce the unobserved heterogeneity. To allow for any dependence, we use the WLW method (Allison, 1995) to get robust variance estimates. This method, however, does not correct for any bias in the coefficients[3].

To account for ties, in which several firms experience the same rating migration event at the same time (t+s), it has been traditional to use approximation adjustments such as the Efron method. However, the SAS software package provides an exact method for handling ties, which we use in this study.

Two generic models for rating migrations were estimated[4]:

- The up-grade model considers up-grade events. In this model, down-grades (rating runs with starting ratings better than ending ratings) and Not rated (NR) runs are considered censored runs.
- The down-grade model considers down-grade events. In this model, up-grades (rating runs with ending ratings better than starting ratings) and NR runs are considered censored runs.

  Many of the current credit pricing models focus on accurately modeling default, which is a special case of down-grades. The consequences of a default are qualitatively and quantitatively different from other down-grades and therefore, we also develop an estimated model for default runs. In this model, up-grade, non-default down- grade, and NR runs are considered censored.

3.3. Variables

Duration, the length of time in years of a rating run, is the dependent variable in the models. As the aim of this research is to investigate how rating migration probabilities depend on rating history, the following independent variables were included in the models:

---

[3] See Allison (1995, Chapter 8)

[4] An all-run model would make no sense since different models with different signs on common variables result for up runs and down runs.

- Lag one: The duration (in years) of the non-censored run immediately preceding the current run (LAG_ONE)
- Lag two: The duration (in years) of the non-censored run immediately preceding lag one run (LAG_TWO)
- The rate of prior rating changes: This rate equals the number of prior rating changes observed between the entry of the firm to the study and the beginning of the current run divided by the time the firm spent in the study (RATE_PRIOR_CHANGE)
  - If the first rated date was after the beginning of this study (1 January, 1986), the time a firm spent in the study is the length in years from the first rated date to the beginning of the current run
  - If the first rated date was before 1 January, 1986, the time a firm spent in the study is the length in years from 1 January, 1986 to the beginning of the current run
- Original rating: the rating of the firm when it was first rated (ORIGINAL_RATING)
- Entry rating: The rating of the firm when it entered the study (1 January, 1986) or as of the date the firm was first rated after the beginning of the study (ENTRY_RATING)
- Start rating: The rating at the beginning of each rating run analyzed (START_RATING)
- Entry 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 rating run (ENTRY_AGE_SINCE_FIRST_RATED)
- A prior Not rated (NR) status: A dummy variable was created to indicate whether the firm underwent a NR status from the time it had entered the study until the beginning of the current rating run (DUMMY_NR)

- Direction of the lag one run: Two dummy variables were created to indicate whether the lag one is a down-grade run (DUMMY_LAG_DOWN) or an up-grade run (DUMMY_LAG_UP). In the estimation sample, no rating run has a censored lag rating one. In other words, they all have either a lag one down-grade run or a lag one up-grade run.

Unlike most studies on rating dynamics, which just focus on the coarser rating categories (AAA, AA), we employ finer rating sub-categories such as AAA, AAA-, AA+, AA- . In our study, a rating run might indicate a finer rating upgrade from A to A+ or a finer rating downgrade from A to A-. For each rating run, we obtain the numeric values of the above covariates.

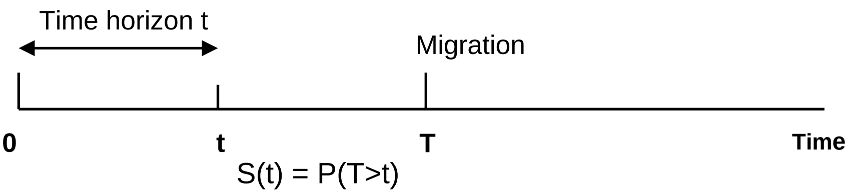
The rating scales, which take into account plus and minus signs, are coded from 0 to 26 with 0 indicating the default state (D) and 26 indicating the AAA state. Details of the rating coding are provided in Table 1. The higher the value of a rating variable (original rating, entry rating, start rating), the better the quality of the firm at the corresponding time. A similar coding technique was employed by Kim and Wu (2006) to examine the impacts of sovereign credit rating history on international capital inflows to emerging countries, and on the development of financial sectors in these countries. 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 number of degrees of freedom and would also hinder compact presentation of the results.

Dummy variables were created to account for the country and the sector of each firm. The sector categories, represented by dummy variables in the model, were the thirteen categories provided by Standard & Poor's in the CreditPro 2005 dataset. A detailed variable description is provided in Table 1.

TABLE 1 HERE

3.3. Forecast technique

The cumulative survival probability measures the probability that a rating run will continue beyond the time horizon t, and is defined as: S(t) = P(T>t), where T is the time that a migration (either a downgrade or an upgrade) occurs.



The estimated survival function S(t) for run i has the form:

$$\hat{S}_i(t) = [\hat{S}_0(t)]^{\underset{\sim}{Z}'\underset{\sim}{\hat{\beta}}}$$

where $\hat{S}_0(t)$ is the estimated baseline of survival function, $\underset{\sim}{\hat{\beta}}$ is the coefficient vector estimated from the Cox model, $\underset{\sim}{Z}'$ is the actual covariates vector.

We use the estimated baseline of survival function and the estimated coefficient vector from in-sample analysis, and use the actual covariates vector from the holdout sample to estimate the survival probabilities for the holdout firms. In this manner, survival estimates, with five-year horizon, are obtained for up-grades, down-grades, and default runs.

In order to form the cumulative accuracy profile (CAP), it was necessary to pick survival probabilities across runs observed at a particular time t. The value for t that we chose was the time at which we identify the run with the survival probability below but closest to 50%. We then categorize each holdout sample into 100 groups in order of ascending survival probabilities**.** The firm with the lowest survival probabilistic estimate is the most risky with respect to the migration event relevant to each model. For instance, in the up-grade model the firm with the lowest survival estimate is the one most at risk of being upgraded. Each holdout sample is then categorized into 100 groups, with 10 firms in a group, in order of descending migration risks. The actual migration status of each firm (run) in each group is recorded, and mapped against the predictions obtained at the survival time t mentioned above. The fraction of total migrations for each group is calculated by dividing the number of migrations in that group by the total migrations observed in the holdout sample. The results are plotted as a CAP curve as in Figure 2.

The CAP curve represents the cumulative fractions of total migrations over 100 groups starting from the most risky group. The curve would indicate a perfect foresight if all migrated runs (firms) are assigned to the groups with the lowest survival probabilistic estimates. Thus, the closer the model's CAP curve to the left of the chart, the more accurate the predictive forecasts produced by the model.

4. Data

Rating data was obtained from Standard & Poor's CreditPro 2005. The whole dataset includes the rating history of 11,605 firms (of which 63.2% are America) over the period 1981-2005. We chose 1986 as the starting point of our study as the high yield bond market in the US was being established as a substantial market in the first half of 1980s. As we focus on rating migrations, which are more common events in the high yield bond sector, migrations consequent to the establishment of the high yield bond market from the middle of the 1980s constitute an important source of events for the study.

The rating runs for each firm in the dataset were identified and those runs were then pooled for all firms. A random estimation sample of 3000 rating runs was selected within the 15 years from 1 January, 1986 to 31 December, 2000. A random holdout sample of 1000 rating runs was selected within the 5 years from 1 January 2001 to 31 December 2005.

- Rating runs that had started prior to 1 January, 1986 entered the estimation sample on 1 January, 1986 with censored durations. Similarly, rating runs that had started prior to 1 January, 2001 entered the holdout sample on 1 January, 2001 with censored durations.
- Rating runs that ended after 31 December, 2000 left the estimation sample on 31 December, 2000 as right censored. In the holdout sample, rating

runs that have not experieneced migrations by 31 December, 2005 are treated as right censored as of 31 December, 2005.

- Rating runs which become unrated (NR) sometime within the 15 years of the study, or during the holdout period are treated as censored runs at the time they left their rating classes

Histograms of run lengths in estimation samples are depicted in Figure 1. Up-grade runs have a non-normal distribution with the heaviest concentration of probability centered at run lengths ranging from 0.25 to 3.75 years. Down-grade runs and default runs have the heaviest concentration of probability at the shorter run lengths.

FIGURE 1 HERE

Statistics for run lengths in the up-grade, down-grade, and default estimation samples are given in Table 2.

TABLE 2 HERE

- In the up-grade estimation sample, 714 runs experience migrations (up-grades) and 2286 runs (76.2%) are censored (including down-grades). Up-grade runs vary from 2 days to 10.44 years, have mean length of 2.19 years, and a median length of 1.71 years.
- In the down-grade estimation sample, 1119 runs experience migrations (down-grades), and 1881 runs (62.7%) are censored (including up-grades). Down-grade runs vary from 1 day to 10.55 years, have mean length of 1.44 years, and a median length of 0.849 years.
- In the default estimation sample, 191 runs experience migrations (defaults) and 2809 runs (93.63%) are censored. Default runs vary from 1 day to 4.08 years, have mean length of 117.27 days and a median length of 60 days.

5. Results

5.1. Model estimation

The following model with 33 variables was estimated for each type of run during 1986-2000:

$$r(t)=r_0(t)\exp\{\beta_1\ \text{Lag\_one}+\beta_2\ \text{Lag\_two}+\beta_3\ \text{Rate\_prior\_change}+\beta_4\ \text{Original\_rating}+\beta_5\ \text{Entry\_rating}+\beta_6\ \text{Start\_rating}+\beta_7\ \text{Entry\_age\_since\_first\_rated}+\beta_8\ \text{Dummy\_NR}+\beta_9\ \text{Dummy\_lag\_down}+\sum_{k=10}^{21}\beta_k D_{\text{sector\_k}}+\sum_{j=22}^{33}\beta_j D_{\text{country\_j}}\}$$

where

$r(t)$ is the hazard of rating migration at time t

$r_o(t)$ is an unspecified baseline hazard function when all variables equal 0

D_{sector_k} is the dummy variable for sector k. $D_{sector\_k} = 1$ for sector k, and $D_{sector\_k} = 0$ for sectors other than k

$D_{country\_j}$ is the dummy variable for country j. $D_{country\_j} = 1$ for country j, and $D_{country\_j} = 0$ for countries other than j.

The above model was estimated for the up-grades, down-grades, and default runs (firms). The results of the estimated models are presented in Table 3.

TABLE 3 HERE

In interpreting Table 3, a negative coefficient reduces the hazard and therefore reduces the probability of a rating migration. The reported hazard ratios (risk ratios) represent the relative change in the hazard for a one unit change in the independent variable. For example, in the up-grade model, an increase in the length of the lag one rating run by one year reduces the chance of an upgrade by (1-0.886) or 11.4%.

The results show that the hazard of a rating change strongly depends upon the rating history prior to the current rating run. The country and sector impacts vary according to the type of run being modeled. Their impact is particularly important in the default model. For example, it is clear that being in Mexico dramatically increases the risk of default. However, there is an issue in relation to the country dummy variables. Their disparate sample sizes naturally lead to questions about the representatives of the samples for future population of rated firms in the different countries.

5.2. Generic model for up-grades and down-grades

The impacts of statistically significant variables on the hazard of a rating migration were as follows:

- Lag one duration (Lag_one): These models show that the longer the length of the lag one rating run, the lower the hazard of migration. An additional year increase in the duration of the lagged one rating run reduces the migration hazard by 11.4% and 8.3% for up-grade runs and down-grade runs respectively. This confirms our expectation of a negative relation between the duration of the lagged one rating run and migration hazard.
- Lag two duration (Lag_two): The up-grade model shows that the longer the lag two rating run, the lower the hazard of migrations. An additional year increase in the duration of the lag two rating run reduces the migration hazard by 8.8% for up-grade runs. This confirms our expectation of a negative relation, and the weaker impact of the duration of the lag two rating run (compared with the length of the lag one rating run) on migration hazard.
- Rate of prior rating change (Rate_prior_change): The down-grade model show that the higher the rate of prior rating changes, the higher the hazard of rating migrations. If the rate of prior rating change increase by one unit, the hazard of migrations increases by 9.8% for down-grade runs. Consistent with our

expectation, a positive relation exists between the rate of prior rating changes and the migration hazard.

- Start rating at run entry (Start_rating): Both models show that the higher the rating at run entry, the lower the hazard of rating migrations. Each additional unit increase in the start rating (for example, an improvement from B- to B) reduces the hazard for up-grades and down-grades by 15.9% and 9.4% respectively. This is consistent with our expectation of the start rating's negative impact on the migration hazard.
- Original rating (Original_rating): The down run model shows that the higher the rating of the firm when it was first rated, the lower the hazard of migrations. Each additional unit increase in original rating reduces the migration hazard by 6.5% for down-grade runs. This is consistent with start rating's impacts. Compared with start ratings, original ratings do not have as strong an impact on migration hazard.
- Entry rating (Entry_rating): Unlike start rating and original rating, this variable has a positive impact on migration hazard for down runs. The higher the rating of the firm when it entered the study, the higher the hazard of a rating change for down-grade runs. Each additional unit increase in entry rating increases the migration hazard by 5.5% for down-grade runs.
- Entry age since first rated (Entry_age_since_first_rated): The up-grade model shows that the older the firm since it was first rated, the higher the migration hazard. An additional year increase in the rating age increases the migration hazard by 4% for up-grade runs. Contrary to our expectation, up-grades rather than down-grades exhibit aging effects. Compared with a newly rated firm, a seasoned firm is more likely to experience an upgrade.
- A prior Not rated (NR) status (Dummy_NR): Both up and down-grade models indicate that if the firm underwent a NR status during the time it has been in the study, this reduced the migration hazard of subsequent runs. The NR status would reduce the migration hazard by 67.9% and 65% for up-grade runs and down-grade runs respectively.
- The direction of the immediately prior rating run (Dummy_lag_down): Consistent with our expectation, a lag one down run would reduce the migration

hazard for up-grade runs and increase the migration hazard for down-grade runs. The lag one down run would reduce the hazard for up-grade runs by 40.1% and increase the hazard by 142.2% for down-grade runs.

- Country dummy: Country variables were significant in the model for up-grade runs only. Being an Australian or a French company increases the hazard of a rating migration by 171.3% and 177% respectively. Germany and Argentina' firms experience few upward migrations, thus the hazard ratio of Country_Germany and Country_Argentina dummies in the up-grade model approaches zero.
- Sector dummy: Sectors variables were significant in the generic models for down-grade runs. Being in the Energy and Natural Resources, Forest and Building Products/ Homebuilders, and Utility sectors lower the migration hazard by 55%, 45.4%, 54.1% for down-grade runs respectively.

5.3.. Model for migrations to default

With regard to rating history variables, the default model has some similarity to both the up-grade and down-grade models. For example, the coefficient of dummy_lag_down has a positive sign for both the down-grade and default models, while the rate_prior_change, original_rating, and entry_rating are insignificant in both the up-grade and default models.

The default model sees statistically significant impacts of such Countries as Mexico, Italy, Germany, Holland, France. Being a firm from Mexico would increase the default hazard while being a firm from the other countries mentioned above would reduce default hazard. Nine out of twelve sectors exhibit statistically significant positive impacts on default hazard.

In summary, the estimated models indicate that rating migration hazards strongly depend on rating history prior to the current rating run. The duration of the lag one rating run, start rating, and a prior Not rated (NR) status have consistent negative impacts on the migration hazards in all models.

5.4. Cumulative Accuracy Profile (CAP)

The maximum run length observed in up-grade and down-grade estimation samples is more than 10 years, but the holdout samples only include rating runs from 1 January, 2001 to 31 December, 2005. The holdout data, therefore, constrains the assessment of predictive accuracy to a five-year horizon.

Using the estimated models and the covariate data from the holdout samples, survival estimates were estimated for each firm in the holdout sample. The accuracy of the estimated survival probabilities are assessed using the CAP curve at a particular time t. We identify time t by locating the run with the survival probability below but closest to 50%. The time t of up-grade, down-grade, and default runs are 2.62 years, 7.1 months, and 2.499 months respectively. The CAP curve of each model is compared with the two benchmark CAP curves:

- A "naïve" CAP curve where rating scores are assigned randomly
- A perfect CAP curve where all cases of rating migrations are assigned to the lowest rating scores.
    - In the up-grade model, the perfect CAP curve would assign the total 156 up-grades (15.6% of the holdout sample) to 16 groups with the lowest survival probabilistic estimates
    - In the down-grade model, the perfect CAP curve would assign the total 382 down-grades (38.2% of the holdout sample) to 39 groups with the lowest rating scores
    - In the default model, the perfect CAP curve would assign the total 96 defaults (9.6% of the holdout sample) to 10 groups with the lowest rating scores

Figure 2 depicts the CAP curves for up-grade, down-grade, and default models.

FIGURE 2 HERE

If “A” denotes the area between the model’s CAP curve and the naïve curve, “B” denotes the area between the naïve curve and the perfect CAP curve, the accuracy ratio of “A” to “B” indicates the predictive accuracy of the model. The accuracy ratio varies from 0 to 100%[5], the higher the accuracy ratio, the better the model.

The CAP curves illustrate that while the up-grade model is just slightly better than the naïve model, the down-grade and default models exhibit good forecast performance.

7. Conclusion

In this study, Cox’s proportional hazard model is employed to estimate time-varying rating migration probabilities. This study examined a number of non-Markovian behaviors in rating dynamics, and assessed the predictive accuracy of probabilistic rating forecasts. Using a sample from Standard & Poor’s CreditPro 2005 dataset, we develop three hazard models for rating up-grades, down-grades, and defaults during 1986-2000 period.

While the significance and sign of the variables for countries and sectors are not consistent across models, we find that there is statistically significant dependence of time-varying migration hazard on rating history prior to the current run.

For both rating up-grades and down-grades, the length of the lag one rating run, the rating at the beginning of the rating run analysed, and a prior Not rated status have negative impacts on the migration hazards. The length of the lag two rating run has a negative impact on the migration hazard for up-grades. We find positive impacts of rating age on up-grade and default hazards, and positive impact of the rate of prior rating changes on down-grade hazard. Of particular interest is the evidence of rating momentum. The lag one down run has a significant negative impact on the migration hazard for up-grade runs, but a significant positive impact on the migration hazard for down-grade runs. Such dependence of migration hazards on rating history suggests that

[5] Computationally, it is convenient to compute the accuracy ratio from the area under the Receiver Operating Characteristic (ROC) curve. This is done using the formula: AR = 2AUC -1 (see Englemann, Hagel and Tasche, 2003). Where: AR is the accuracy ratio, AUC is the area under the ROC curve.

rating migrations follow a non – Markovian process. These results call into question the Markov assumption common in the analysis of rating migrations.

The Cumulative Accuracy Profile (CAP) curve is used to assess the predictive accuracy of probabilistic rating forecasts. While the up-grade model provides only a slightly better forecast ability than a random forecast, the down-grade and the default models exhibit good accuracy in assigning the lowest survival estimates to the riskiest firms. Given the asymmetric impacts of rating down-grade announcements on stock abnormal returns[6], the predictive ability of the down-grade and default models appears worthy of further development.

[6] See Atlman and Kao (1992), Carty and Fons (1994), Lando and Skodeberg (2002), Hamilton and Cantor (2004), Figlewski, Frydman and Liang (2006)

References

Allison, P. D. (1984). “Event History Analysis”, Sage University Paper Series on Quantitative Applications in the Social Sciences.

Allison, P. D. (1995). “Survival analysis using SAS A practical guide”, SAS Press

Altman, E. (1998). “The importance and subtlety of credit rating migration.” Journal of Banking and Finance 22: 1231-1247.

Altman, E., and D. L. Kao (1992). “Rating Drift of High Yield Bonds.” Journal of Fixed Income March: 15-20.

Altman, E., and H. Rijken (2005). “The Impact of the Rating Agencies’ Through-the-cycle Methodology on Rating Dynamics.” Economic Notes by Banca Monte dei Paschi di Scienna SpA 34: 127-154.

Avramov, D., G. Jostova, A. Philipov (2003). “Corporate Credit Risk Changes: Common Factors and Firm-Level Fundamentals.”

Berd, A. (2005). “Dynamic Estimation of Credit Rating Transition Probabilities.”

Cantor, R., and C. Mann (2003). “Measuring the performance of corporate bond ratings.” Moody’s Investor Service Special Comment, April.

Carty, L.V., and J. S. Fons (1993), “Measuring changes in corporate credit quality”, Moody’s Special Report, New York.

Cantor, R., J. Fons, C. Mahoney, D. Watson, and K. Pinkes (1999). “The evolving meaning of Moody’s bond ratings.” Moody’s Investor Service Rating Methodology August.

Choy, E., S. Gray, V. Ragunathan (2006). “Effect of credit rating changes on Australian stock returns.” Journal of Accounting and Finance 46: 755-769.

Cox, D. (1972). “Regression Models and Life Tables.” Journal of Royal Statistical Society Series B (Methodological) 34: 187-220.

Cox D., a. D. O. (1984). Analysis of Survival Data, Chapman & Hall, London.

Engelmann, B., E. Hayden, and D. Tasche (2003). “Measuring the discriminative power of rating systems.” Discussion Paper, Deutsche Bundesbank.

Figlewski, S., H. Frydman, W. Liang (2006). “Modeling the Effects of Macroeconomic Factors on Corporate Default and Credit Rating Transitions.”

Frydman, H., and T. Schuermann (2006). “Credit ratings dynamics and Markov mixture models.” Working paper.

Gonzalez, F., F. Haas, R. Johannes, M. Persson, L. Toledo, R. Violi, M. Wieland, and C. Zins (2004). “Market dynamics associated with credit ratings: a literature review.” Occasional paper series (European Central Bank) 16.

Hamilton, D., P. Varma, S. Ou, and R. Cantor (2006). “Default and recovery rates of corporate bond issuers, 1920-2005.” Moody’s Investor Service Special Comment, March.

Hamilton, D., R. Cantor (2004). “Rating transitions and default conditional on watchlist, outlook and rating history.” Moody’s Investor Service Special Comment, February.

Helwege, J., and P. Kleiman (1996). “Understanding Aggregate Default Rates of High Yield Bonds.” Federal Reserve Bank of New York Current Issues in Economics and Finance 2: 1-6.

Hu, J., and R. Cantor (2003). “Structured finance rating transitions: 1983-2002.” Journal of Portfolio Management.

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.”

Lando, D (2004) “Credit Risk Modeling”, Princeton University Press

Lando, D., and T. Skodeberg (2002). “Analyzing ratings transitions and rating drift with continuous observations.” Journal of Banking and Finance 26: 423-444.

Mann, C., D. Hamilton, P. Varma, and R. Cantor (2003). “What happens to fallen angels? A Statistical review 1982-2003.” Moody’s Investor Service Special comment, July.

Nickell, P., W. Perraudin, and S. Varotto (2000). “Stability of ratings transitions.” Journal of Banking and Finance 24: 203-222.

Partington, G., and M. Stevenson (2001). “The probability and timing of price reversals in the property market.” Journal of Applied Econometrics 18: 23-46.

Santarelli, E. (2000). “The duration of new firms in banking: an application of Cox regression analysis.” Empirical Economics 25: 481-499.

Yao, Y., G. Partington, and M. Stevenson (2005). “Run length and the predictability of stock price reversals.” Journal of Accounting and Finance 45 (December, 2005): 653-671

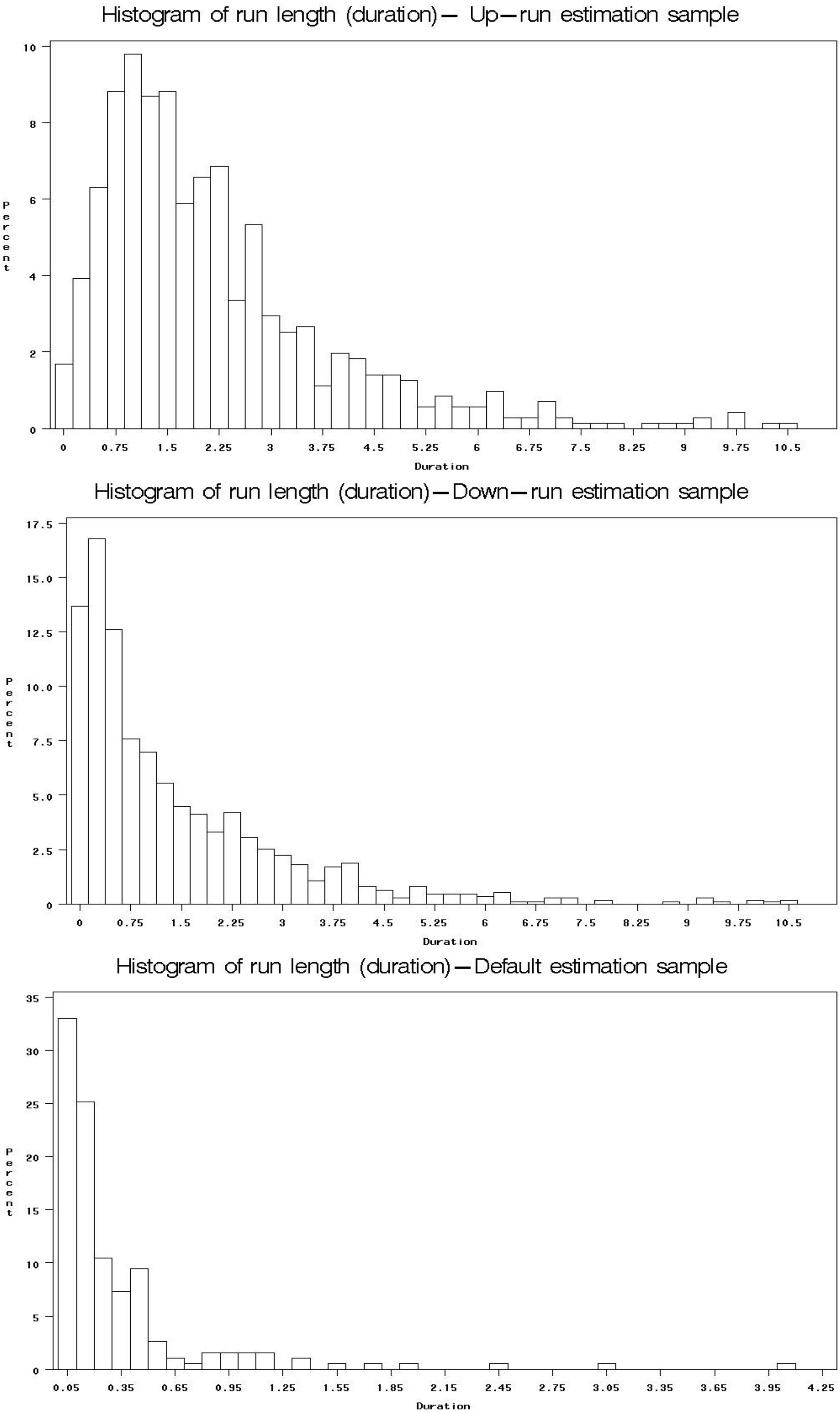
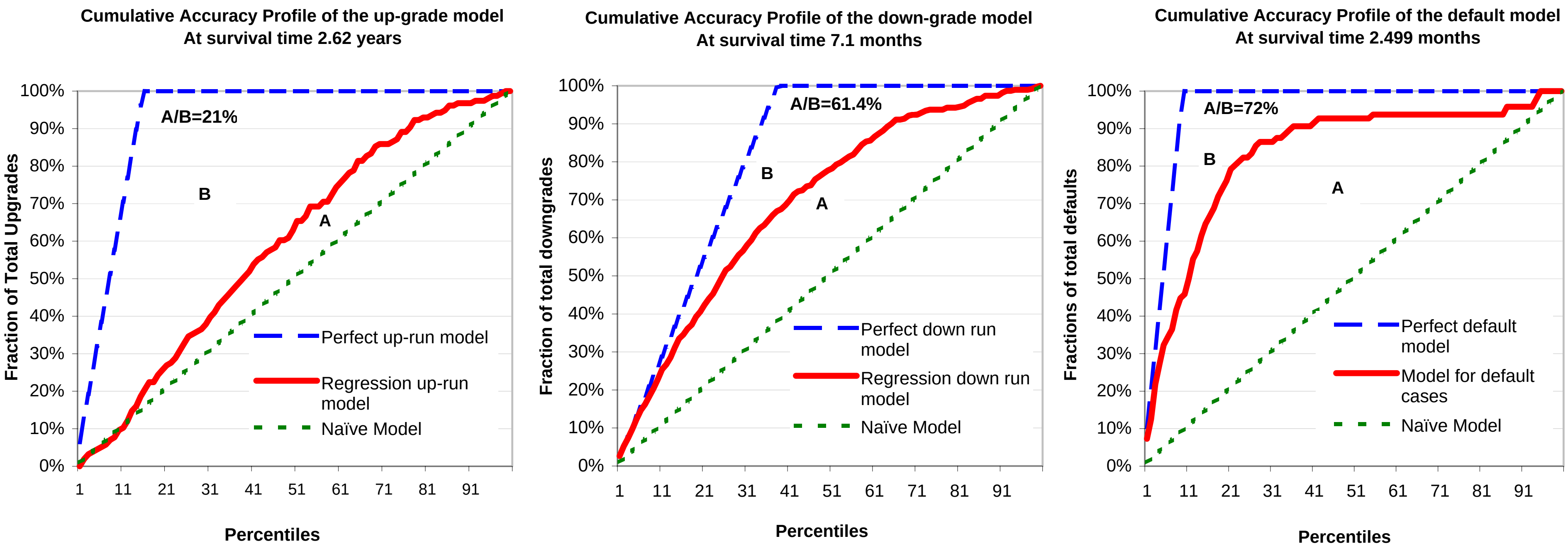


Fig. 1: Histogram of run length (Duration).



Up-grade holdout sample		
Censored runs	844	84.40%
Up runs	156	15.60%

Down-grade holdout sample		
Censored runs	618	61.80%
Down runs	382	38.20%

Default holdout sample		
Censored runs	904	90.40%
Default runs	96	9.60%

A naïve predictive model would assign migrated runs equally throughout the chart, that is, the weakest 10% of firms (runs) correspond to 10% of runs with lowest survival probability estimates and so on. The CAP curve of the naïve model is the diagonal straight line. "A" denotes the area between the CAP curve of the regression model and the CAP curve of the naïve model.

A perfect predictive model would assign all migrated firms (runs) in the weakest firms, which have the lowest survival probability estimates. "B" denotes the area between the CAP curve of the perfect predictive model and the the CAP curve of the naïve model.

The more predictive the regression model, the more the model's CAP curve cluster to the left of the chart into the weakest firms (runs) groups. The accuracy ratio of the area "A" to area "B" measures the predictive accuracy of the regression model. This ratio varies from 0 (no predictive ability) to 100% (perfect predictive ability). This ratio is 72%, 61.4%, and 21% for the default, down-grade, and up-grade models respectively.

Fig. 2 : Cumulative Accuracy Profile .

Table 1
Variable dictionary

Variable (in bold letters)	Description	Codes/ Values
First_rated_date	Date the firm was first rated	
Start_date / End_date	The starting date / ending date of each rating run	
Duration (Dependent variable)	The length of a rating run	Years (End date - Start date) / 365
Entry_age_since_first_rated	Rating age (since it was first rated) at run entry	Years (Start_date -First_rated_date) / 365
Start_rating	The starting rating at the beginning of each rating run	0=D 3=C+ 7=CCC- 11=B 15=BB+ 19=A- 23=AA
Entry_rating	The rating at the time the firm entered the study	1=C- 4=CC- 8=CCC 12=B+ 16=BBB- 20=A 24=AA+
Original_rating	The orginial rating when the firm was first rated	2=C 5=CC 9=CCC+ 13=BB- 17=BBB 21=A+ 25=AAA-
		6=CC+ 10=B- 14=BB 18=BBB+ 22=AA- 26=AAA
Lag_one	The length of the non-censored lag one run	Years
Lag_two	The length of the non-censored lag two run	Years
Rate_prior_change	Rate of prior rating changes measured from either the beginning of the study* or since the firm was first rated (whichever is later)	(365 * Number of prior rating changes) / Days spent in the study
Dummy_NR	Dummy variable indicating whether the firm underwent a NR status during the time it spent in the study	
Dummy_lag_down	Dummy variable indicating whether the immediate prior rating run was a down run or not	
Dummy_lag_up	Dummy variable indicating whether the immediate prior rating run was an up run or not	
Sector **	<u>**Firm's sector coded as a dummy variable**</u>	**Financial institutions**
	Aerospace / automotive / capital goods / metal	**Insurance**; **Forest and building products / homebuilders**
	Consumer / service sector	**Leisure time / media**; **Health care / chemicals**
	Energy and natural resources	**Real Estate**; **Transportation**
	Telecommunications	**Utility**; **High technology/ computers/ office equipment**
Country (whole dataset)	<u>**Firm's country coded as a dummy variable**</u>	**Holland (112 firms)** 0.97%; **Japan (410 firms)** 3.53%
Total: 11605 firms ***	**US (7330 firms)** 63.16%	**France (257 firms)** 2.21%; **Italy (132 firms)** 1.14%
	Canada (401 firms) 3.46%	Taiwan (140 firms) 1.21%; **Mexico (202 firms)** 1.74%
	UK (421 firms) 3.63%	**Argentina (105 firms)** 0.90%; **Germany (225 firms)** 1.94%
	Australia (304 firms) 2.62%	**Brazil (136 firms)** 1.17%; Others 12.32%

* 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. 12 Sector Dummies (without Sec_Telecommunications) were included in the model

*** The whole CreditPro 2005 dataset includes the rating history of 11,605 firms from 13 "active" countries specified above, each of which represents about 1% or more of the total firms in the whole dataset, and a group of the remaining countries (others) . The 3000 firm estimation sample (1986-2000) does not have rating data related to Taiwanese firms. Thus, 12 dummy variables (without Country_Other and Country_Taiwan) were included in the model.

Table 2
Descriptive statistics of duration (run length)

Model	Number of non-censored runs	Non-censored runs/ sample size* (percentage)	Mean (years)	Median (years)	Stdandard Deviation	Skewness	Kurtosis	Min (days)	Max (years)
Default runs	191	6.37%	0.3213	0.1644	0.502	4.089	22.245	1 day	4.0795
Down-runs	1119	37.30%	1.443	0.849	1.656	2.121	5.947	1 day	10.55
Up-runs	714	23.80%	2.19	1.708	1.758	1.725	3.667	2 days	10.44

[*] sample size includes 3000 runs during 1986-2000

Table 3:
Summary of significant variables in regression models during 1986-2000

Significant variables***	Up-run model (76.2% censored)			Down-run model (62.7% censored)			Default Model (93.63% censored)		
	Parameter estimate	Standard error	Hazard ratio	Parameter estimate	Standard error	Hazard ratio	Parameter estimate	Standard error	Hazard ratio
Lag_one	-0.12086*	0.02667	0.886	-0.08668*	0.02439	0.917	-0.26159**	0.10252	0.77
Lag_two	-0.09181*	0.02972	0.912						
Rate_prior_change				0.09339*	0.02576	1.098			
Original_rating				-0.06708*	0.02186	0.935			
Entry_rating				0.05307**	0.02653	1.055			
Start_rating	-0.17311*	0.01694	0.841	-0.09914*	0.01573	0.906	-0.56112*	0.03125	0.571
Entry_age_ since_ first_ rated	0.03922*	0.01161	1.04				0.06033**	0.02859	1.062
Dummy_NR	-1.13632*	0.13923	0.321	-1.05109*	0.12262	0.35	-1.70470*	0.40818	0.182
Dummy_lag_down	-0.51297*	0.08914	0.599	0.88475*	0.08172	2.422	1.64315*	0.51454	5.171
Country_Australia	0.99817*	0.33472	2.713						
Country_Germany	-10.03058*	0.61733	<0.001						
Country_France	1.01876*	0.36672	2.77						
Country_Argentina	-11.50529*	0.58697	<0.001						
Country_Mexico							1.59026*	0.57005	4.905
Country_Italy							-9.34425*	0.76624	<0.001
Country_Germany							-9.11430*	0.9099	<0.001
Country_Holland							-10.22285*	0.72126	<0.001
Country_France							-9.14690*	0.66397	<0.001
Sec_Aerospace_automotive							2.66357**	1.18737	14.347
Sec_Consumer_Service							2.58233**	1.1683	13.228
Sec_Energy_natural_resources				-0.79778*	0.28144	0.45			
Sec_Forest_building_products				-0.60521**	0.2604	0.546	2.84395**	1.18949	17.184
Sec_Financial_Institutions							2.40695**	1.17614	11.1
Sec_Healthcare_chemicals							2.55690**	1.19914	12.896
Sec_Leisure_time_media							2.57797**	1.16696	13.17
Sec_Transportation							2.93533**	1.20243	18.828
Sec_Utility				-0.77972*	0.22963	0.459	2.86713**	1.22167	17.586

* $p \leq 1\%$ based on Wald Chi-square tests.

** $1\% < p \leq 5\%$ based on Wald Chi-square tests

*** The sample during 1986-2000 does not have rating runs related to Taiwan. Thus, only 12 Country dummies, and 12 sector dummies were included in the models. Insignificant country dummies and sector dummies were not presented