# Issuer-heterogeneity and time-heterogeneity in the rating migration dynamics of U.S. financial institutions


## Abstract

This study examines issuer-heterogeneity and time-heterogeneity in the rating migration dynamics of U.S. financial institutions during the period 1984-2006. The study found that (i) different downgrade outcomes require separate models while upgrade outcomes can be treated equivalent in the same analysis; (ii) rating history, macro-economic and political conditions are the key determinants of a subsequent rating change. These effects persist after controlling for the current rating, the outlook/ CreditWatch designation; (iii) during the holdout period (2007-2010) issuer-heterogeneity and time-heterogeneity jointly exhibit good forecast performance for upgrades and downgrades to high investment ratings, and show some calibration power for downgrades to mid and low ratings; (iv) accounting for the outlook/ CreditWatch listing, in most cases, does not improve the predictive accuracy of the models out-of-sample. The findings explicitly rule out the Markov and time-homogeneity properties inherent in the discrete time cohort Markov framework, which is commonly used by credit rating agencies to estimate a rating migration matrix.



**Keywords:** financial institutions, rating migrations, issuer-heterogeneity, time-heterogeneity, hazard model, forecast accuracy, Brier score, Murphy decomposition

**Issuer-heterogeneity and time-heterogeneity in the rating migration dynamics of U.S. financial institutions**

## 1. Introduction

Since the Basel II framework came into effect, credit ratings have been widely used by banks to assess counterparty credit risks and to determine capital adequacy requirements. A concern for any bank is to adopt an appropriate approach in modelling rating migration probabilities of its counterparties. A small change in a probability estimate may result in a substantial variation in regulatory capital requirement. Tsaig, Levy and Wang (2011) indicated that credit migration can explain as much as 51 percent of volatility and 35 percent of economic risk capital for a typical loan portfolio. Jafry and Schuermann (2004) suggested that changing the estimation method of rating migrations leads to more variation in economic risk capital than switching between economic contraction and expansion.

In practice the discrete time cohort Markov framework has been widely used by credit rating agencies to construct a rating migration matrix. This modelling framework assumes that the current rating alone determines the probability of a future rating re-grade (Markov property) and the rating migration process is static across time (time-homogeneity property). However, literature on corporate rating dynamics suggests that the Markov property does not persist at a horizon longer than one or two years (Kiefer and Larson, 2007; Frydman and Schuermann, 2008). Issuers of the same rating grade migrate at different rates and the heterogeneity exists after controlling for business cycles and industry sectors (Frydman and Schuermann, 2008). The source of within-rating heterogeneity can be attributed to a variety of aspects of rating history[1]. There is also strong evidence that rating stability varies over time and ratings move pro-cyclically[2]. Failing to consider heterogeneity in rating migration process can result in inaccurate estimates of credit risk and a misleading picture of the economic risk capital (Nickell, Perraudin and Varotto, 2000; Kadam and Lenk, 2008).

Since the onset of the global financial crisis credit rating agencies have been increasingly criticised for their excessive tendency to assign FIs investment grade ratings

---

[1] See, for example, Altman and Kao (1992a, 1992b), Carty and Fons (1994), Nickell et al. (2000), Kavvathas (2001), Bangia et al. (2002), Lando and Skodeberg (2002), Hamilton and Cantor (2004), Kadam and Lenk (2008), Figlewski, Frydman and Liang (2012)

[2] See Altman and Kao (1991), Blume, Lim, MacKinlay (1998), Nickell, Perraudin and Varotto (2000), Kavvathas (2001), Bangia, Diebold, Kronimus, Schagen, and Schuermann (2002), Koopman, Lucas, and Monteiro (2006), Bolton, Freixas and Shapiro (2009), Figlewski, Frydman and Liang (2012)

which understated their inherent risks[3]. Deficient credit rating agencies' models have been identified as a cause of the global financial crisis (Hellwig, 2008; Hull, 2009). The need to develop a robust rating migration model that accurately captures the credit quality of FIs and that has a predictive power for future rating changes has been emphasised since then. Such modelling requires a thorough understanding of the rating behaviours of FIs. Literature has documented that issuers in the financial sector exhibited different rating dynamics compared to issuers in other sectors (Nickell et al., 2000; Lando and Skodeberg, 2002; Kadam and Lenk, 2008)[4]. However, the empirical analysis of issuer-heterogeneity and time-heterogeneity in FIs' rating process has been somewhat piecemeal. This study attempts to fill this gap and focuses on U.S. FIs. Inclusion of FIs domiciled in countries other than the U.S. was considered, but sample sizes across countries were often small and were highly variable over time. The study is thus restricted to U.S. data.

This study aims to answer the following questions: (i) Does one migration model fit all, or do different migration outcomes require different models? (ii) How do within-rating heterogeneity and time-heterogeneity affect subsequent re-grades?; Are these effects robust in the presence of the current rating, the outlook or CreditWatch designation?; (iii) Do within-rating heterogeneity and time-heterogeneity exhibit forecast performance out of sample?; and Does their forecast performance change after controlling for the current rating, the outlook or CreditWatch listing?

The study, employing a rich issuer rating dataset of FIs in the U.S., extends the literature as follows. First, new evidence is offered on issuer-dependence and time-dependence in the rating migration dynamics of FIs during the estimation period 1984-2006. This study is the first attempt to examine the effects of a number of past rating behaviours and of the U.S. political environment on the migration dynamics of FIs. Some aspects of rating history have not been addressed in previous studies, for example, the sequence of the current rating in a firm's rating history, the average magnitude of prior rating changes, the average survival time a FI stayed in a rating grade, the history of a frequent fallen angel or rising star and large rating change events. The political cycle is also highly relevant as FIs are

[3] Rating agencies depend on their reputation capital and if such capital fluctuates pro-cyclically, they may have an incentive to set ratings pro-cyclicality (Ferri, Liu, Stiglitz, 1999, p.352). Bolton *et al.* (2009) showed that the reputation risk of credit rating agencies understating credit risk is lower during economic expansions, which gives the agencies an incentive to assign ratings that understate risk during economic booms.

[4] For example, Lando and Skodeberg (2002) separately considered the rating observations of issuers in the financial sector. They found that the duration effect and downward momentum are less pronounced for FIs than for other obligors in other sectors.

particularly vulnerable to changes in fiscal and monetary policies, which typically occur following the presidential and/ or congress elections. Second, the study contributes to the framework for modelling rating migrations by developing a robust empirical model that overcomes some limitations of the conventional discrete time cohort Markov approach. The proposed Cox's dynamic hazard model accounts for repeated migrations of the same issuer, accommodates both time-fixed and time-varying covariates in the estimation process, and permits a rigorous testing of issuer-heterogeneity and time-heterogeneity in rating dynamics. This study is the first attempt to employ outlook and CreditWatch listing as time-varying covariates, thereby capturing the deterioration in the credit quality of a FI while it retains a rating grade. Third, there is increasing interest in estimating survival probabilities of FIs that are dynamic in nature and have predictive power. The study overcomes the computational challenges involved in forming time-varying probability survival estimates when the proportional assumption of the conventional Cox's hazard model (Cox, 1972) does not hold[5]. The dynamic forecasts may aid regulators in the early detection of financially impaired FIs and in developing an early warning system as an effective off-site monitoring tool[6]. Fourth, the study presents new evidence on the predictive ability of within-rating heterogeneity and time-heterogeneity in estimating time-varying survival probability for holdout rating observations during the global financial crisis, January 2007-September 2010.

The study found that different downgrade outcomes require separate models as their hazards are neither equal nor proportional as time passes. Upgrades outcomes, in contrast, can be treated equivalent in the same analysis. For each major migration outcome, several models were estimated to examine the effects of issuer- and time-heterogeneity in the rating dynamics. Overall, the migration hazard of FIs depends significantly upon several aspects of rating history, macro-economic factors and the political business cycle. Different migration outcomes are characterised by different significant factors. Downgrades to investment rating boundary and downgrades to speculative ratings are more impacted by past rating behaviours whereas downgrades to high investment ratings and upgrades are more sensitive to adverse macro-economic and political conditions. The observed issuer-heterogeneity and time-heterogeneity persists after controlling for the outlook/ CreditWatch status.

[5] Figlewski et al (2012) developed dynamic Cox's hazard models to investigate the effects of time-varying macro-economic variables on the hazard of rating migrations and defaults. They did not construct time-varying probability estimates and therefore were unable to assess the forecast performance of the models out of sample.

[6] It is suggested that Cox's proportional hazard model "identified failed and healthy banks with a high degree of accuracy", and "flagged a large proportion of banks that subsequently failed as potential failures in periods prior to their actual demise" (Whalen, 1991, p. 21) .

The current rating state has a modest impact in comparison to some aspects of past rating behaviours, for example rating momentum. The downward momentum is more pronounced the lower the destination ratings and is stronger in the presence of the CreditWatch. For example, a downgrade at lag one rating makes a plunge to junk grade ratings 61.6 times more likely in the presence of the current rating but 784.1 times more likely after controlling for the CreditWatch status. The outlook is not significant in determining the probability of any downgrade outcomes whereas the CreditWatch placement has a strong effect on the hazard of a downgrade to high/ mid investment ratings.

During the holdout period January 2007-September 2010, past rating behaviours, macro-economic and political cycle jointly exhibit predictive ability, particularly for downgrades to high investment ratings which dominate the holdout sample. The forecast performance of these factors is in direct contrast to the suggestion that the Markov property adequately holds within a one or two-year horizon (Kiefer and Larson, 2007; Frydman and Schuermann, 2008). Adding the CreditWatch listing, in most cases, does not improve the performance of the estimated models out of sample.

The remainder of this paper is structured as follows. Section 2 reviews the literature, Section 3 describes the data. Section 4 presents the competing risks analysis. Section 5 presents the estimation method, Section 6 summarises the estimation results. Section 7 presents the forecast method and forecast performance assessment. Section 8 summarises the key findings, limitations, and implications of the study.

## 2. Literature review

### 2.1. Outlook and CreditWatch listings

This study employs Standard & Poor’s issuer rating data. The following literature review therefore focuses on its outlook and CreditWatch placement process. According to Standard & Poor’s (2009), its issuer ratings reflect a predictive view about the creditworthiness of issuers  and take into account future events to the extent they can be reasonably anticipated. Outlook and CreditWatch listings address the possibility that future performance differs from initial expectations, with a focus on the scenarios that could lead to a rating change. There is strong empirical evidence that CreditWatch listings and outlook status exhibit forecast performance for future rating changes[7].

[7] See  Hamilton and Cantor (2004),Vazza, Leung, Alsati, and Katz (2005), Hill, Brooks, and Faff (2010), Bannier and Hirsch (2010), Guttler (2011), Al-Sakka and Gwilym (2012)

A CreditWatch listing indicates the potential change to a short or long term rating. Standard & Poor's places a rating on CreditWatch if there is at least a one-in-two likelihood of a rating change within the next 90 days. If a rating remains on CreditWatch for more than 90 days or if material events or deviations from trends occur, Standard & Poor's generally publish interim updates to reflect its current assessment of the situation. If ratings are placed on CreditWatch due to performance deterioration of securitized assets or due to a change in criteria, and the analysis shows the impact is expected to exceed 90 days, Standard & Poor's will generally publish a timeframe during which it expects to complete its assessment. CreditWatch designations may be positive or negative. Occasionally, Standard & Poor's may assign a developing CreditWatch in situations when future events are so uncertain that the rating could be raised, lowered, or affirmed. A CreditWatch listing does not mean a rating change is inevitable and does not imply that any potential change would be only one notch (Standard & Poor's, 2009). Bannier and Hirsch (2010) examined Moody's corporate issuer ratings and found that for the high quality borrowers, the watchlist procedure is mainly used to deliver precise and stable information in order to satisfy investors' demands. On the other hand, for low quality issuers the watchlist is used as an implicit contract in order to induce the rated companies to abstain from further risk enhancing actions.

Rating outlook indicates the potential direction of a long term credit rating over the intermediate term. Outlooks have a longer time horizon than CreditWatch listings and incorporate trends or risks that Standard & Poor's believe have less certain implications for credit quality. However, the potential for change must be realistic and not remote. Standard & Poor's assign positive or negative outlooks to issuer ratings (except when the rating is on CreditWatch) when it believes that an event or trend has at least a one-in-three likelihood of resulting in a rating action over the intermediate term for investment grade credits (generally up to two years) and over the shorter term for speculative grade credits (generally up to one year). Occasionally where the outlook refers to a longer timeframe, Standard & Poor's will explicitly say so in its published analysis. Standard & Poor's may assign a developing outlook to an issuer when it believes that a rating may be raised or lowered. A positive or negative outlook is not necessary a precursor of a rating change or a CreditWatch listing. Conversely, rating changes can occur when the issuer has a stable outlook. If warranted, the rating would be changed to reflect the most current opinion of credit quality, and Standard & Poor's would not delay such a change by revising the outlook or placing a rating on CreditWatch, merely to signal a potential change (Standard & Poor's, 2009).

## 2.2. Issuer-heterogeneity in rating migration dynamics

The current rating grade affects rating stability and future rating distribution[8]. As credit quality declines volatility of rating migrations increase sharply (Nickell et al., 2000). Furthermore, issuers rated around the investment and speculative grade boundary (BBB-, BB+) exhibit different propensities compared to their peers in neighbouring rating categories (Carty and Fons,1994; Carty, 1997; Johnson, 2004).

There is strong evidence that rating migrations display downward momentum. A downgrade is more likely to be followed by a further downgrade [9], and a one notch rating change in low rating grades implies a larger increase in default risk (Jorion and Zhang, 2007). The effect of a lagged rating change, however, becomes weaker with the passage of time and diminishes once the rating outlook or the CreditWatch status is controlled for (Hamilton and Cantor, 2004). Fledelius, Lando, and Nielsen (2004) suggested that the downward momentum lasts no longer than two or three years.

Issuers with a lagged rating change of large magnitude are more likely to experience a future re-grade[10]. There is also evidence of mean reversion for issuers of mid-rating grades or low end-investment grades. Issuers initially rated in the middle of the rating universe do not exhibit a tendency to substantially drift in either direction, and ratings tend to migrate toward the middle of the rating spectrum. (Altman and Kao, 1992b; Kavvathas, 2001).

Subsequent rating changes also depend on the duration of a rating state (Carty and Fons, 1994; Lando and Skodeberg, 2002). Issuers staying a short period in sub-investment grades tend to experience high re-grades and default (Koopman et al., 2006). Lando and Skodeberg (2002) further suggested that the evidence of duration dependence and downward momentum are not as strong for FIs as for issuers in other sectors.

Issuers with different original ratings show different migration dynamics and retain their original ratings in different ways[11]. The time since an issuer was first rated also affects future rating distribution (Altman, 1998, pp. 1239-1240). Issuers of new bonds generally have lower credit risk and may retain their original ratings for a longer period of time than their peers with seasoned bonds.

[8] See, for example, Lucas and Lonski (1992), Carty and Fons (1994), Carty (1997), Hamilton and Cantor (2004), Jorion, Shi, and Zhang (2009), Figlewski et al. (2012).
[9] See Altman and Kao (1992a, 1992b), Carty and Fons (1994), Kavvathas (2001), Bangia et al. (2002), Lando and Skodeberg (2002), Hamilton and Cantor (2004), Figlewski et al. (2012)
[10] See Guttler and Wahrenburg (2007), Al-Sakka and Gwilym (2009), Bannier and Hirsch (2010)
[11] See Altman and Kao (1991, 1992a, 1992b), Jorion et al. (2009), Figlewski et al. (2012).

Different rating paths lead to different rating distributions. Fallen angels who crossed the investment grade boundary (BBB-/BB+) are riskier than their peers following their fall date (Mann, Hamilton, Varma, and Cantor, 2003; Vazza, Aurora, and Schneck, 2005). They exhibit strong downward momentum, experience a rapid migration rate until reaching their lowest rating grades, and are vulnerable to default. However, over extended periods, fallen angels possess robust franchise value, enhanced business strength and improved profitability. Compared with their peers, they exhibit a greater tendency to survive and to rebound strongly to investment grades after surviving the initial years of financial distress.

The above evidence suggest that within-rating heterogeneity can be attributed to different aspects of rating history such as the direction of the lagged rating state, lagged rating duration, the magnitude of lagged rating change, mean rating, the incidence that the current rating is higher than the mean rating, rating age, the original rating, and a history of fallen angel events. This emphasises the need to account for the past rating behaviours described above in estimating future rating re-grades. This study also extends the literature by exploring additional aspects of rating history which has received little attention, such as the sequence of the current rating, a history of substantial rating changes, a history of frequent rising star events, the mean magnitude of prior rating changes, the mean survival time, the occurrence that the magnitude of lag one rating change is larger than the mean magnitude, the incidence that the lag one rating duration is longer than the mean survival duration, and rating volatility.

### 2.3. Time-heterogeneity in rating migration dynamics

Rating migrations are principally affected by macro-economic factors rather than the characteristics of debt issues (Blume, Keim, and Patel, 1991). There is strong evidence that corporate rating dynamics differ in times of recession and growth (Bangia et al., 2002). Downgrades and defaults occur more often during periods of contraction whereas upgrades occur more often during periods of growth. Rating volatility decreases during business cycle peaks and increases during troughs (Nickell et al., 2000). Low ratings are more vulnerable to adverse macro conditions than high ratings. According to Bangia et al. (2002, p. 469), failure to incorporate macro-economic factors in credit risk models may lead to an underestimation of “downward potential of high yield portfolio” in contractions or “suboptimal capital allocation in lending business.”

The political cycle in the U.S., particularly the presidential and congress elections, may be another source of time-heterogeneity in rating behaviours. The Democratic Party, in contrast to the Republican Party, does not believe in the power of the "invisible hand" and holds the view that the government plays an important role in regulating the behaviour of market participants. For example, Progressive Democrats tend to support higher taxes on wealth, distrust Wall Street, favour more government support to prevent mortgage foreclosures and support the nationalisation of troubled banks[12]. The term of a Democratic President, or periods when the President's Party lost seat(s) in the mid-term Congress election and becomes the minority in the Congress tends to be associated with frequent or unexpected changes in fiscal and monetary policies, which may affect FIs' business environment unfavourably. Besides, it is suggested that incumbent governments tend to manipulate fiscal and monetary policies to encourage voter support during the calling of national elections[13]. Block and Vaaler (2004) observed an increase in sovereign rating downgrades during the years when national elections occur. This emphasises the need to examine several aspects of the political cycle, for example, being in a presidential election year, being in the term of a Democratic Party's President, being in the year when the President's Party lost seat(s) in the mid-term Congress election, or when the President's Party is not the dominant Party in the Congress.

## 3. Data

### 3.1. Estimation and holdout periods

The study covers a long period from January 1984 to September 2010 and thereby includes several business cycles in the U.S. economy. The estimation period 1 January 1984[14]-31 December 2006 saw two economic recessions (July 1990-March 1991, March 2001-November 2001), the U.S. stock market crash in 1987, the Mexican currency crisis in 1994, the Asian financial crisis in 1997, the Russian sovereign bond default in 1998, the collapse of the Long-Term Capital Market Hedge Fund in 1998, the dot-com bubble burst in 2000, the devastating 9/11 terrorist attack in 2001, the U.S. bond crisis in 2002-2003 and the dramatic bankruptcies of fallen angels like WorldCom and Enron. The period after the estimation period, January 2007 - September 2010 was used to construct a holdout sample for model validation purpose. This period witnessed the sub-prime mortgage crisis in 2007-2009,

[12] The Progressive Democrats of America consider themselves a grassroots Political Action Committee operating inside the Democratic Party. They support government-centred programs and oppose policies that offer free-market solutions (See www.pdamerica.org).
[13] See, for example, Beck (1987), Grier (1989), Haynes and Stone (1989, 1990, 1994), Klein (1996)
[14] 1984 was chosen as the year of commencement for two reasons. First, Standard & Poor's rating scales were changed in 1983. Second, macro-economic data was not all available prior to 1982 and the macro-economic variables used were constructed in the form of 18 months of distributed lags.

a prolonged economic recession from December 2007 to June 2009 and the unprecedented bankruptcies of a number of investment-grade rated FIs.

### 3.2. Corporate issuer rating data

The corporate issuer rating data of U.S. FIs were obtained from Standard & Poor's Ratings Xpress on 28 September 2010. For each issuer in the data I extract the full sequence of rating history, outlooks and CreditWatch listings. The exact migration dates between rating classes, varying from AAA to D (default), are recorded. These dates are then used to determine the durations in each rating grade. Since the data on the last rating change is required for some of the rating history variables, only firms experiencing at least one prior migration and having CreditWatch and outlook designations for at least one day during the study period were included in the final dataset. A coding approach often used in the literature was adopted to replace Standard & Poor's alphabetical rating scales by numeric scales, varying from 0 to 21 with 0 being the default state (D) and 21 representing AAA rating[15].

The estimation and the holdout samples include 447 and 114 rating observations respectively. Of the 447 estimation states, 130 (29.08 per cent) experienced downgrades and 153 (34.23 per cent) experienced upgrades. Of the 114 holdout observations, downgrades and upgrades respectively contribute 67 (58.78 percent) and 11 (9.65 percent) states. The holdout period showed a substantial increase in downgrade frequency and a marked decline in upgrade frequency, reflecting a sharp deterioration in the credit quality of FIs. Additional statistics (not reported) show that the proportions of downgrades (upgrades) in the estimation and holdout period are statistically different.

### 3.3. Time to events (survival time)

A rating state begins when a FI enters a rating grade (*start rating*) after the start of the study and ends when the FI migrates to another grade (*ending rating*), withdraws from being rated, or the study period ends. The survival time of each rating state is the time a FI maintains the same grade.

The histogram of time to upgrades and time to downgrades (survival time) for rating states in the estimation period 1984-2006 are depicted in Figure 1. Upgrade and downgrade

[15] This coding approach has been widely adopted (Sy, 2002; Kim and Wu, 2008; Al-Sakka and Gwilym, 2009; Hill, Brooks, Faff, 2010). The numeric rating scales maintain the rank order of the alphabetical scales, capture fine revisions intra-rating, and allow for a compact presentation of the results. Besides, two dummy variables employed in the models control for any non-linearity in the rating scales surrounding the boundary between investment and speculative grades (BBB-/BBB/BBB+, BB-/BB/BB+).

states both showed positively skewed distributions. Downgrade states had a markedly shorter survival time than upgrade states and are heavily massed in durations shorter than a year.

FIGURE 1 HERE

The descriptive statistics of the survival time for rating states in the study are given in Table 1. Additional analysis (not reported) indicates that downgrade (upgrade) states in the estimation and holdout periods had statistically different survival times. It took substantially much less time for a migration to occur in the holdout period than it did in the estimation period.

TABLE 1 HERE

## 4. A competing risks analysis of alternative migration outcomes

There are multiple ending rating grades that a FI could migrate to. However, a firm cannot simultaneously be in more than one rating grade at any point in time. The occurrence of one migration outcome removes a firm from the risk of all other migration outcomes at that point in time. Thus, different downgrades (upgrades) to different destination ratings *may* be treated as competing risks. Due to the small sample size and sparse migration events to finer rating grades, attempting to estimate rating transitions between individual grades would result in low statistical power and make the assessment of the model's predictive power for rare migration events difficult. Alternatively, pooling all downgrades (upgrades) to the current rating state as down states (up states) allows a larger number of observations in the at-risk population. However, lumping different migration events into the same analysis clouds the results for specific outcomes and results in misspecification. An estimation approach that accommodates migrations to broad destination ratings overcomes the issue of small sample size/ sparse events and provides an opportunity to investigate the impacts of within-rating heterogeneity and time heterogeneity across major migration paths.

The observed frequencies of downgrades (upgrades) categorised by broad destination ratings are presented in Table 2 Panel A. Downgrades/ upgrades to A- or higher ratings respectively account for 53.1 per cent and 74.3 per cent of the total downgrades and upgrades observed in the estimation period. This is not surprising given that FIs generally have high credit quality; most of them were assigned investment grade ratings during the study period.

TABLE 2 PANEL A HERE

Empirical studies show that upgrades do not result in much, if any, market response. By contrast downgrades are consistently associated with a statistically significant negative return in both the equity and bond markets. Price reaction is more significant for downgrades crossing the boundary between investment and speculative grades, and for downgrades within the speculative grade category[16]. In common with the literature, this study hypothesises that downgrades to high/ mid investment ratings and downgrades to investment rating boundary (BBB-, BBB, BBB+) are likely to have quite different determinants than downgrades to speculative ratings. I had no such prior expectations with regards to alternative upgrade outcomes.

To lay the foundation for the case that different migration routes follow different processes it is necessary to conduct a formal statistic test for hazard proportionality. The relevant question is if the hazard for a downgrade (upgrade) outcome changes with time, will the hazard for other downgrade (upgrade) paths change by a proportional amount? To answer this question, the following multinomial logit model was estimated (Cox and Oakes, 1984).

$$Log[h_j(t)] = \alpha_0(t) + \alpha_j + \beta_j t \qquad (1)$$

Where:

$Log[h_j(t)]$ is the logarithm of the hazard for the contrast between two migration types $j$ at time $t$. For this analysis, downgrades to speculative ratings (BB+ or lower), which mostly represent fallen angels[17], were contrasted with other downgrade outcomes while upgrades to investment rating boundary (BBB-/BBB/BBB+), which primary represent rising stars, were contrasted with other upgrade events.

For downgrades, $j$ =1, 2, 3 respectively, indicates the contrast between a downgrade to AA-/AA/AA+, a downgrade to A-/A/A+, a downgrade to investment rating boundary (BBB-, BBB, BBB+) versus a downgrade to speculative ratings (BB+ or lower).

For upgrades, $j$ =1, 2, 3, respectively, indicates the contrast between an upgrade to speculative ratings (BB+ or lower), an upgrade to mid-investment ratings (A-, A, A+), an upgrade to high

[16] Holthausen, and Leftwich, 1986; Moody's, 1994; Hite and Warga, 1997; Goh and Ederington, 1999; Jorion and Zhang, 2007

[17] Regulations either do not allow institutional investors to hold speculative grade securities or require that extra capital be held against these securities (Dale and Thomas, 1991; Cantor and Packer, 1997). When an issuer loses its investment grade status, the pool of potential investors is significantly reduced. Such fallen angels are no longer in demand from all investors but will have to rely on the small pool of investors to which the aforementioned restrictions do not apply (Ferri, Liu and Stiglitz, 1999, p. 335-336). A soaring cost of capital and significant liquidity issues tend to beset these fallen angels, particularly if multiple triggers occur simultaneously (Standard & Poor's, 2001).

investment rating (AA- or higher) versus an upgrade to investment rating boundary (BBB-, BBB, BBB+).

$t$ is the survival time (i.e. event time) a firm stayed in a rating grade. Observations that did not experience a downgrade (an upgrade) during the study period were excluded from the analysis for downgrades (upgrades).

$\beta_j$ is the survival time coefficient for the contrast between the two event types $j$

The results of the models estimated as per equation (1) for 130 downgrades and for 153 upgrades are presented in Table 2 Panel B and Panel C. Panel B provides the analysis of variance output. Panel C presents the coefficient estimates derived from the maximum likelihood procedure.

TABLE 2 PANEL B AND PANEL C HERE

Under the proportional hazards hypothesis, the coefficient for survival time $t$, $\beta_j$, will be zero. As shown in Panel B, the effect of survival time $t$ for downgrades is significant at 1 per cent level, indicating a rejection of the proportional hazards hypothesis. The parameter estimates for survival time $t$ (Panel C) show that the $\beta$ coefficients for the contrasts between downgrades outcomes are statistically significant, and their log-hazards diverge non-linearly with time. Compared with downgrades to speculative ratings, other downgrade events have hazard rates that increase more rapidly with time[18]. For upgrades, the effect of survival time $t$ (Panel B) and the $\beta$ coefficients for the contrast between major outcomes (Panel C) are not statistically significant. Thus, the proportional hazards hypothesis for observed upgrade paths cannot be rejected.

The above analysis suggests that downgrades to alternative broad destination ratings follow different processes which govern both the occurrence and the timing of the events. Thus, alternative downgrade outcomes should be modelled separately as competing risks while alternative upgrade events can be treated equivalent and pooled in the same analysis[19].

## 5. Estimation method

### 5.1. Estimation model

[18] The log-hazard ratio for the contrast $j$=1, 2, 3 between downgrades to AA-/AA/AA+, downgrades to A-/A/A+, downgrades to BBB-/BBB/BBB+ versus downgrades to BB+ or lower, respectively, increases by $100(e^{1.2171}-1)=24.25\%$; $100(e^{1.5011}-1)=65.05\%$; $100(e^{1.4274}-1)=53.33\%$ each year.

[19] This is also convenient for the assessment of the upgrade models' forecast performance given rare upgrade events observed in the holdout sample

This study adopts the survival analysis framework (Allison, 1984) and develops Cox's hazard model with time-varying covariates to model the hazards of different migration outcomes [20]. The following dynamic Cox's hazard model was estimated for generic upgrades and different downgrade outcomes:

$$h_{s,m}(t, Z, Z(t)) \;=\; h_{s(0)}(t)\;\exp[Z_{\mathrm{j}}^{\mathrm{m}}\beta_{\mathrm{j}} + Z_{p}^{m}(\mathrm{t})\beta_{\mathrm{p}}\;] \qquad (2)$$

Where: $h_{s\,m}(t, Z, Z(t))$ is state $m$'s hazard of a migration outcome $s$ at time $t$ given its time-independent covariate vector $Z_j^m$ and its time-varying covariate vector $Z_p^m(t)$. $h_{s(0)}(t)$ is the baseline hazard of a migration outcome $s$ at time $t$. $\beta_p$ is the vector of estimated coefficients for time-varying covariates $Z_p^m(t)$. $\beta_j$ is the vector of estimated coefficients for time-independent covariates $Z_j^m$.

Rating observations are arranged in event time (gap time) risk sets, which are composed of all rating states that are at risk of an event at time $t$. A new risk set is formed at each time $t$ when a migration event of interest occurs. In forming the risk sets for an outcome-specific downgrade hazard model, the downgrade events of that type (outcome) are treated as events. In forming the risk sets for the generic upgrade model, all upgrade outcomes are treated as equivalent events. If a FI leaves the study due to any other reason apart from the migration event of interest, its survival time is treated as censored. Rating states ending after the observation period are also regarded as censored.

A FI may contribute several rating states to the dataset, which may lead to dependence among observations. This problem is minimised in two ways. First, the covariates in the models control for the sequence of rating changes and for the dependence among rating states of the same issuer. Second, the study uses the marginal-event specific method proposed by Wei, Lin, Weissfeld (1989) to account for dependence among rating states of the same issuer.

Appendix A present in details the estimation of the time fixed and time-varying covariates, the estimation of the baseline hazard function, and the generation of time-varying probability survival forecasts for holdout rating states.

[20] The advantages of using survival analysis framework to model rating migrations have been well articulated by Kavvathas (2001), Lando and Skodeberg (2002), Guttler (2011).

### 5.2. Covariates

The candidate variables that capture issuer-heterogeneity and time-heterogeneity in rating migration dynamics were identified from the literature. I used a more extensive set of variables than in prior literature, and some variables I used have not been explored in previous studies. The list of time-fixed and time-varying variables employed in this study and their definition are presented in Table 3. Time-independent variables were measured at the beginning of a rating state whereas the value of a time-varying variable used in the estimation process was updated to the most recent value as a migration event of interest occurred.

TABLE 3 HERE

#### 5.2.1. Firm-specific variables

Three time independent variables were used to control for the current rating state and its proximity to the boundary between investment and speculative grade (*start rating*, *dummy investment boundary* and *dummy junk boundary*). The two dummy variables capture any non-linear effects surrounding the investment and speculative rating threshold[21].

Two time-varying variables were created to control for the positive and negative outlook designation, and to indicate the potential direction of the long term credit rating over the intermediate term (typically six months to two years) (*dummy positive outlook*, *dummy negative outlook).* Two time-varying variables were also included to account for the positive and negative CreditWatch listing, and indicate the likelihood of rating action within the next 90 days (*dummy positive CreditWatch*, *dummy negative CreditWatch).* The values of these dummies were updated monthly during the survival time of each rating observation.

The rating history variables utilise the full rating history of each FI over the study period. Seventeen time-independent variables were employed to capture various aspects of past rating behaviours such as the sequence of the current rating in a firm's rating history (*rating sequence*), the first rating received *(original rating),* the direction of lagged one rating change (*dummy lag one down,)* the duration of lagged one rating state *(lag one duration),* the magnitude of lagged one rating change (*lag one rating change magnitude),* the average magnitude of prior rating changes (*mean rating change magnitude),* the average rating received since a firm was first rated (*mean rating),* the average time a firm stayed in a rating grade (*mean survival time),* the incidence that the magnitude of lagged one rating change is

[21] Moody's (1994) reports that yields are relatively unresponsive to downgrades when ratings remain in the investment grade territory but become very sensitive to even small downgrades when ratings plunge to the speculative grade spectrum.

larger than the average magnitude of prior rating changes (*dummy lag one magnitude > mean rating change magnitude*), the occurrence that the duration of lagged one rating is longer than the average survival time (*dummy lag one duration > mean survival time*), the incidence that the current rating is better than the average rating (*dummy start rating > mean rating*), upgrade and downgrade volatility *(rate prior upgrades, rate prior downgrades),* the number of prior fallen angel/ rising star events (*number fallen angel events, number rising star events),* and the number of prior substantial rating jumps (*number big downgrade events, number big upgrade events).*

In addition, one time-varying variable was constructed to capture the rating age of each firm (*Logarithm of age since first rated*). Rating age (*Age since first rated)* is a function of the time since a firm was the first rated. Linear effects from increasing rating age as rating duration increases are automatically absorbed into the baseline hazard (Hosmer, Lemeshow and May, 2008). Thus, for rating observation *m* I use a nonlinear function of age and survival time *t* as follows:

$$Logarithm\ of\ \ age\ since\ first\ rated_t^m = Log\ \ (Age\_since\_first\_rated^m + t\ ) \qquad (3)$$

The estimation process requires an updated value of survival time *t* in equation (3) whenever a migration event of interest occurs.

### 5.2.2. Macro-economic and political cycle variables

Four time-varying variables were constructed to account for the political business cycle in the U.S. *(dummy presidential election year, dummy Democratic party's President, dummy President's party lost seat in mid-term congress election, dummy President party's not the dominant party in the Congress*).

Ten time-varying variables were employed to account for the U.S. macro-economic conditions[22]. *Dummy NBER recession* captures the recession state of the U.S. economy. *Capacity utilisation* and *output gap* controls for the general level of U.S. economic activity. *Inflation expectation* signals the future prospects of the economy. *Default spread* and *total U.S. corporate debt defaulting* reflects credit conditions. *SP500 Index Return, SP 500 return standard deviation, and cyclically adjusted PE ratio for the aggregate stock market* represents the performance of the stock market. *FI industry's corporate default* accounts for the credit risk in the financial institution sector.

[22] Twenty eight candidate macro-economic variables were considered and those that showed strong multicolliniarity were eliminated, leaving ten macro-economic variables which were used in the models.

As macro-economic conditions tend to affect the rating dynamics of FIs with a lag, an exponentially weighted average of lagged observations computed monthly over a window of 18 months was applied to construct macro variables other than *Dummy NBER recession*. The construction of macro-economic lagged values is similar to the approach applied by Figlewski et al. (2012).

The values of *Inflation expectation, Capacity utilization, Cyclically adjusted PE ratio for the aggregate stock market, S&P500 Index return, S&P500 return standard deviation,* and *Default spread* were updated monthly whereas the value of *Output gap* was updated quarterly and the values of *FI Industry's corporate default rate* and *Total U.S. corporate debt defaulting* were updated yearly during the survival time of each rating observation. The values of macro-economic and political dummies were updated monthly and entered the estimated models without any transformation.

### 5.3. Statistics of rating and macro-economic variables

The descriptive statistics of rating variables for observations in the estimation and the holdout samples are given in Table 4. Both the estimation and holdout samples showed a dominance of ratings in the investment grade categories as indicated by the mean and median of *start rating*. This is not surprising as the investment grade rating has been the norm in the financial sector. Being confidence and capital sensitive entities, it is difficult for FIs to operate with poor credit quality. Downgrade momentum was more pronounced in the holdout period than in the estimation period as seen by the mean and median of *dummy lag one down*. Additional analysis (not reported) shows that observations in the estimation and the holdout samples have statistically different rating history.

TABLE 4 HERE

Each rating observation may have several CreditWatch/ outlook listings during its survival time. The descriptive statistics of CreditWatch (outlook) durations and CreditWatch (outlook) history for rating-CreditWatch (rating-outlook) observations are presented in Table 5. Of 447 rating states in the estimation period, there were 686 rating-CreditWatch observations and 1157 rating-outlook observations. Among 686 rating-CreditWatch observations, 166 were with a positive CreditWatch and 161 were with a negative CreditWatch. Of 1157 rating-outlook observations, 135 were with a positive outlook and 209 were with a negative outlook. Rating-outlook observations with a positive (negative) outlook had a longer mean duration than rating-CreditWatch observations with a positive (negative)

CreditWatch. Furthermore, rating-CreditWatch (rating-outlook) observations with a positive/ negative CreditWatch (outlook) experienced a greater number of prior positive/ negative CreditWatch (outlook) designations. This suggests a momentum in CreditWatch (outlook) designations, a fact highlighted by Al-Sakka and Gwilym (2012).

TABLE 5 HERE

The descriptive statistics of the time series for the exponentially weighted average of macro-economic variables used in this study are shown in Table 6. Additional analysis (not reported) indicates that the macro-economic conditions in the estimation and the holdout periods were statistically different. The holdout period observed a sharp decline in *output gap*, a negative stock return (*S&P500 Index return)*, a substantial increase in total corporate debt default *(Total U.S. corporate debt default in US$ billion)* and a deterioration in the credit quality of financial institutions as indicated by the industry's higher default rate (*FI industry's corporate default rate)*

TABLE 6 HERE

## 6. Estimation Results

The estimation results of 3 models for 3 downgrade outcomes and 1 model for generic upgrade events are given in Table 7 Panel A. Table 7 Panel B provides the statistics on the goodness of fit of the estimated models. Table 7 Panel C summarizes the number of upgrade/ downgrade events and censored observations in the estimation and holdout samples.

TABLE 7 HERE

The effects of issuer heterogeneity and time heterogeneity on upgrades, downgrades to A- or higher ratings[23], downgrades to BBB-/BBB/BBB+, and downgrades to BB+ or lower ratings were first examined in the absence of the current rating (Model 1). Model 2 examines

[23] Owing to its low observed frequency and sparse events in the estimation period, downgrades to high investment ratings (AA-/AA/AA+) were merged with downgrades to mid-investment ratings (A-/A/A+) in the same analysis. To account for any heterogeneity in the baseline hazards of the two downgrade routes, the models for the combined outcome (i.e. downgrades to A- or higher ratings) were estimated with the hazard function being stratified by the destination rating categories *v* (A-/A/A+, AA-/AA/AA+) as follow:

$$h_{s,m,v}(t,Z,Z(t)) \;=\; h_{s,v(0)}(t)\ \exp[Z_{\mathrm{j}}^{\mathrm{m}}\beta_{\mathrm{j}}+Z_{p}^{m}(\mathrm{t})\beta_{\mathrm{p}}\ ]$$

Where: $h\, s_{,m,v}(t, Z, Z(t))$ is the hazard of a downgrade outcome *s (*event *s*, in this case, includes all downgrades to A- or higher ratings) for observation *m* in strata *v* at time *t* given its time-independent covariate vector $Z_j^m$ and its time-varying covariate vector $Z_p^m(t)$. $h\, s,v_{(0)}(t)$ is the baseline hazard of a downgrades to A- or higher ratings for all rating states in strata *v* at time *t.* $\beta_p$ is the vector of estimated coefficients for time-varying covariates $Z_p^m(t)$. $\beta_j$ is the vector of estimated coefficients for time-independent covariates $Z_j^m$.

whether the heterogeneities found in Model 1 persist after taking into account the current rating state. Model 3 and Model 4 respectively examines whether issuer-dependence and time-dependence properties persist after controlling for the outlook/ CreditWatch listing[24].

In estimating a parsimonious model the backward stepwise estimation procedure was used. Variables were retained according to the log-likelihood ratio test, at the 10 per cent level or better, derived from the maximum likelihood procedure used to estimate the models. As shown in Table 7 Panel A, most of the retained variables were significant at the 10 percent level or better based on a Wald chi-square test. The following discussion focuses on the key significant variables present in the estimated models for different migration outcomes. For each type of event, focus will be given to the effects that persist after controlling for the current rating (Model 2), the outlook (Model 3), and CreditWatch status (Model 4) [25].

### 6.1. Upgrades

#### 6.1.1. Issuer-heterogeneity in upgrade dynamics

It is consistent across the models for upgrades that two aspects of rating history, rating age (*Logarithm of Age since first rated*) and the direction of the lag raring change (*Dummy lag one down*), are significant in determining the probability of a future upgrade. Older (well-established) FIs are more likely to improve their credit quality whereas those experiencing a downgrade at lag one state exhibit an unfavourable tendency towards upgrades.

*Number big upgrade events* becomes significant as the current rating, the outlook/ CreditWatch variable(s) is added to the model. Issuers with a history of substantial jumps to higher ratings have a higher probability of ascending the rating scales (Models 2-4). This is not surprising as ratings tend to change in a predictable fashion (Hamilton and Cantor, 2004) and if revised, ratings are only partly adjusted (Altman and Rijken, 2004). If a celebrated upgrade merely represents a partial revision to a substantial improvement in the credit quality of an issuer, the jump may raise the probability of a subsequent upgrade.

A frequent downgrade history (*rate of prior downgrades*), in the absence of the CreditWatch status, is significant in determining the hazard of a rating bounce. As shown in Model 1-3, a higher downgrade volatility is associated with a more than 100 per cent higher

[24] An issuer placed on CreditWatch does not carry an outlook during the CreditWatch review period. Therefore the current (or most recent) outlook and CreditWatch s were not considered simultaneously in the same analysis.
[25] For downgrades to speculative ratings (BB+ or lower ratings), the result of Model 3 (with start rating and outlook) is the same as the result of Model 2 (with start rating). Therefore, only Model 1, 2, and 4 are presented for this downgrade outcome.

upgrade probability. Only one aspect of rating history exhibits inconsistent impact across upgrade models, which is *mean rating*. A higher average rating raises the probability the current rating persists (Model 1) but it makes an upgrade more likely as the current rating, the outlook/ CreditWatch listing are taken into consideration (Model 2-4).

The current rating (*start rating*) exhibits a modest impact while the outlook designation (*dummy positive outlook*) and the CreditWatch status (*dummy positive CreditWatch*) have substantial effects on upgrades. FIs with a better current rating tend to retain the rating grade (Model 2-4) while those with a *positive outlook*/ a *positive CreditWatch* placement are, respectively, 2.09 times (Model 3)/ 26.6 times (Model 4) more likely to ascend the rating scale.

#### 6.1.2. Time-heterogeneity in upgrade dynamics

With regards to macro-economic conditions, it is consistent across models that FIs exhibit an unfavourable tendency towards upgrades in periods characterised by a high corporate default rate in the financial sector, a high expected inflation rate, and a volatile stock market (Model 1-4). On the other hand, FIs are more likely to improve their credit quality during an economic recession or in periods characterised by a large *default spread*, a high *capacity utilisation*, and a high *P/E ratio for the aggregate stock market* (Model 1-3). These macro variables, however, are not significant after controlling for the CreditWatch listing. The impacts of *NBER recession* and *default spread* are particularly pronounced. The positive effect of an economic recession is intuitively understandable given the business nature of FIs. Banks and other FIs operate in a highly regulated business environment and are subject to strict controls regarding capital adequacy and loss reserves. Bank lending standards tend to be most lax during economic growth (Lown et al, 2000) and banking supervisors tend to be most vigilant during economic contraction (Syron, 1991). Banks particularly perform well towards the end of a recession when demands for credit and lending activities increase in anticipation of economic recovery. Consequently, they tended to be resilient in economic downturns.

With respect to the political cycle, being in the term of a *Democratic Party's President* or in periods when the *President's Party is not the dominant Party in the Congress* raises the probability the current rating will persist.

### 6.2. Downgrades

#### 6.2.1. Issuer-heterogeneity in downgrade dynamics

As shown in Panel A Table 7, downgrades to low investment ratings (BBB-/BBB/BBB+) and downgrades to speculative ratings (BB+ or lower) show more dependence on rating history than downgrades to high investment ratings (A- or higher). The models for three downgrade outcomes share some common rating variables. The current rating (*start rating*) is significant in all models in which it is considered but its sign differs across downgrade events. A higher *start rating* is associated with a higher risk of a downgrade to investment ratings but it makes a downgrade to junk ratings less likely, which is consistent with the mean reversion property. Issuers generally attain an "average rating" under normal conditions (Kavvathas, 2001, pp. 32-33) and there is a tendency for ratings to migrate toward the middle of the spectrum (Altman and Kao, 1992b, p. 70).

The effect of *start rating* is modest compared with the effect of being rated BBB-/BBB/BBB+ (*dummy investment boundary)* or being rated BB-/BB/BB+ (*dummy speculative boundary*). FIs rated around the investment and speculative grade boundary exhibit an unfavourable tendency towards downgrades. This finding is consistent with Johnson (2004)'s suggestion that the lowest investment grade ratings are more likely to be downgraded than their neighbouring ratings. The effect of being around the investment/ speculative threshold is harsher the lower the destination rating and is more pronounced after controlling for the CreditWatch status. For example, FIs rated around the investment rating boundary are 2.07 times (26.89 times) more likely to descend to BBB-/BBB/BBB+ (BB+ or lower grades) (Model 2). After the CreditWatch listing is considered, these FIs are 7.29 times (55.95 times) more vulnerable to the respective downgrade outcome (Model 4).

There is consistent evidence of rating momentum (*dummy lag one down*) in the dynamics of downgrades to A- or higher and downgrades to speculative ratings. The downward momentum is particularly strong in the latter case, which is consistent with the evidence of rating drift in corporate rating dynamics. However, in contrast to previous studies, it is found that the momentum becomes more pronounced in the model controlling for the *negative CreditWatch* dummy. For example, a downgrade at lag one rating makes a plunge to junk ratings 61.6 times more likely in the presence of the current rating (Model 2) but 784.1 times more likely after controlling for the CreditWatch placement (Model 4).

The models for downgrades to BBB-/BBB/BBB+ and the models for downgrades to speculative ratings (BB+ or lower) feature some significant rating history variables with large coefficients. There are three common rating history variables, which are the *number of fallen angel events*, *lag one rating change magnitude,* and *rate prior upgrades.* The effects of these variables persist after accounting for the outlook/ CreditWatch status (Model 3-4). FIs with a history of fallen angel events (*number of fallen angel events*) and a large *lag one rating change magnitude*, are less susceptible to downgrades to BBB-/BBB/BBB+ but more vulnerable to downgrades to junk ratings. *Number of fallen angel events* has a strong impact on both outcomes. Its effect on downgrades to BBB-/BBB/BBB+ is more pronounced while its impact on downgrades to speculative ratings become weaker in the presence of the CreditWatch listing (Model 4). A frequent upgrade history (*rate prior upgrades*), on the other hand, has the same effect on both downgrade events; it makes a future downgrade to either low investment ratings or speculative ratings more likely.

The outlook designation (*dummy positive outlook, dummy negative outlook)* is not significant in any downgrade models (Model 3) whereas the CreditWatch status (*dummy negative CreditWatch*) is present in all models in which it is considered (Model 4). Of three examined outcomes, downgrades to investment rating boundary and downgrades to A- and higher are substantially impacted by the CreditWatch status. A negative designation makes the former 212.18 times and the latter 5.53 times more likely. Downgrades to junk ratings, on the other hand, are more affected by being rated around the boundary of investment and speculative ratings, as discussed above.

Apart from the common rating behaviours discussed above, there are some differences between the models for three downgrade outcomes. For downgrades to speculative ratings (BB+ or lower), several rating history variables such as the first rating received (*original rating)*, the *mean rating change magnitude*, the occurrence of a prior rising star event (*number rising stars)*, the incidence of a prior large upgrade (*number big upgrades*) were only present in the absence of the current rating and the CreditWatch listing (Model 1). On the other hand, some other aspects of rating history became significant when the CreditWatch status is included. A long *mean survival time* or a longer than average survival time at lag one rating (*Dummy lag one duration* > *mean survival time*) increases the probability the current rating continues (Model 4). In contrast, a long *lag one duration*, a frequent upgrade history (*number prior upgrades)*, or a larger than average rating magnitude at lag one state (*Dummy lag one magnitude* > *mean rating change magnitude*) are associated with a higher hazard of a plunge to speculative grades (Model 4). The effects of the two latter variables were particularly large. FIs with a more frequent upgrade history or with the lag one rating magnitude being larger than the average magnitude are respectively 3.78 times and 9.25 times more vulnerable to deterioration in credit quality.

Unlike other migration outcomes, downgrades to investment rating threshold (BBB-/BBB/BBB+) do not exhibit rating momentum (*dummy lag one down)* when the current rating, the outlook/ CreditWatch status are taken into account (Model 2-4). The models feature some unique significant rating history variables. Issuers with a better *original ratin*g or those with the current rating higher than the average rating are more likely to retain the investment grade status. These effects persist in the presence of the outlook/ CreditWatch designation (Model 3-4). In addition, *rating sequence*, *age since first rated* (in logarithm form) and *mean rating change magnitude* are significant, except when the CreditWatch enters the model. FIs with a higher *rating sequence* or ageing issuers are at a higher risk of downgrades whereas those with a large *mean rating change magnitude* have a higher probability to stay in the current rating (Model 1-3). The presence of the CreditWatch designation, while masking the effects of these variables, highlights the strong effect of *number rising star events.* A history of rising star events substantially raises the downgrade risk by 19.53 times (Model 4).

For downgrades to high investment ratings (A- or higher), the models with the current rating, the outlook/ CreditWatch status feature fewer significant rating history variables than the respective models for the other two downgrade outcomes. FIs with a higher *rating sequence* tend to experience deterioration in credit quality, so do rising stars (*Number rising star events)* or issuers with a large rating change at lag one state (*lag one rating change magnitude*) and a large mean magnitude of prior rating changes (*mean rating change magnitude*). Being a rising star has a large impact, making a downgrade to A- or higher 8.41 times more likely (Model 1). Issuers with a frequent downgrade history (*rate prior downgrades)* or a high *original rating,* on the other hand, are more likely to maintain the current rating grade. The effects of *rating sequence* and *rate prior downgrades* do not persist once the CreditWatch status is considered whereas the other variables (*original rating*, *lag one rating change magnitude*, *mean rating change magnitude*, *number rising star events*) are only significant in the absence of the current rating. *Mean rating* is the only variable that changes its sign across models. After controlling for the CreditWatch listing, a higher *mean rating* makes the current rating more likely to continue (Model 4).

#### 6.2.2. Time-heterogeneity in downgrade dynamics

As shown in Panel A Table 7, FIs at different stages of financial deterioration exhibit different vulnerabilities to the economic and political environment. Downgrades to high investment ratings (A- or higher) are more susceptible to adverse macro-economic conditions and political cycle than downgrades to investment rating boundary and downgrades to speculative ratings. This can partly be attributed to the fact that the investment grade rating has been the norm in the financial sector. Downgrade candidates have strong incentives to maintain the investment grade status. Candidates to high investment grade status tend to be large firms, with high leverage, large trading assets and are highly interconnected in the financial market. The effects of unfavourable economic and political conditions tend to be harsher for these FIs due to the opaqueness and interconnectedness in their business nature.

Controlling for the outlook or CreditWatch status, in several cases, mask the effects of macro-economic and political conditions. This is not surprising as the CreditWatch/ outlook is issued to reflect the short term/ intermediate term credit prospect of an issuer, and this evaluation is largely influenced by the prevailing environment. The outlook/ CreditWatch designation does not look through the cycle and its time-varying placement therefore capture the evolution of the macro-economic and political cycle.

#### 6.2.2.1. Macro-economic conditions

FIs are more susceptible to downgrades to A- or higher ratings in periods with a high corporate default rate in the financial sector or a stock market bubble (*P/E ratio for the aggregate stock market)*. On the other hand, high rated FIs are more likely to retain the current rating in periods characterised by a large *default spread,* high *capacity utilization,* and a stock market boom (*SP500 Index return)*. The effects of *FI sector's corporate default*, *default spread* and *P/E ratio for the aggregate stock market* persist in all models for downgrades to high investment ratings (Model 1-4). *Default spread* has a particularly large impact, a one per cent increase makes this downgrade outcome 99.99 times less likely. *Capacity utilization* only appears in the model with the CreditWatch status (Model 4) whereas *SP500 Index return* is significant in the other three models (Model 1-3).

The models for downgrades to BB+ or lower share with the models for downgrades to A-/ higher (discussed above) two common macro-economic variables, *P/E ratio for the aggregate stock market*, which captures the stock market bubble, and *SP500 Index return*, which captures the stock market performance. For downgrades to junk ratings, the effects of these variables persist in the presence of the current rating and the CreditWatch status (Model 2-4). FIs are more likely to retain their current rating during a stock market boom (*SP500 Index return)* while they exhibit an unfavourable tendency towards this downgrade outcome in times of a high *inflation expectation* or a high *P/E ratio for the aggregate stock market*.

*Inflation expectation* only appears when the CreditWatch status is not considered (Model 1-3).

For downgrades to the investment rating boundary, *Total U.S.'s debt defaulting* is the only variable that is significant in all estimated models (Model 1-4). Periods characterised by a high default volume (in US$ billion) or a large *output gap* observe more downgrades to BBB-/BBB/BBB+ whereas periods with a high *inflation expectation* is associated with a lower risk. It is worth noting that controlling for the CreditWatch status reinforces the effects of macro-economic conditions on this downgrade outcome as evidenced by the presence of *output gap* and *inflation expectation* (Model 4).

#### 6.2.2.2. Political cycle

The political cycle is not significant in determining the hazard of a downgrade to speculative ratings whereas it exhibits a strong impact on the probability of a downgrade to A- or higher ratings. A higher risk of this downgrade outcome is associated with the time when the *President's Party lost seats in the mid-term Congress election*. In contrast, fewer downgrades to high investment ratings are observed during the year when the Presidential election occurs (*Dummy presidential election year*) or in the periods when the *President's Party is not the dominant party in the Congress*. Controlling for the CreditWatch status diminishes the effect of *Dummy President's party lost seat in mid-term congress election* and eliminates the strong impact of *Dummy President's Party not the dominant party in Congress.*

For downgrades to investment rating boundary, the term of a President representing the Democratic Party (*Dummy Democratic Party's President*), the years when the Presidential election occurs (*Dummy presidential election year*) or when the *President's Party lost seat in mid-term Congress election* are associated with a higher hazard of this downgrade outcome. Election years are characterised by uncertainties in the outcome of the election and election-related manipulations of fiscal/ monetary policies. As FIs are particularly vulnerable to changes in fiscal and monetary policies, it is to be expected that credit rating agencies are more likely to revise FIs' ratings downward during an election year. This tendency tends to be more pronounced for FIs who are already candidates to descend to the investment rating boundary. The political cycle variables, however, is not significant in the models with the outlook/ CreditWatch listing.

Overall, the above analysis suggest that several aspects of rating history, macro-economic conditions and the political cycle jointly have a strong impact on the probability of a future rating change. While the models for each migration outcome somewhat differs, it is clear that the issuer-heterogeneity and time heterogeneity are present in FIs' rating migration dynamics even after controlling for the current rating, the outlook/ CreditWatch status. These key determinants can therefore be used to forecast future migrations, but how accurate are such forecasts?

## 7. Predictive forecast assessment

### 7.1. Forecast method

The Brier score (Brier, 1950) is used to assess the forecast performance of the estimated models on a holdout sample of 114 rating states pooled over the subsequent period, January 2007- September 2010. The Brier score is the average squared error difference between the estimated survival probability and the actual survival outcome of holdout rating observations. The Brier score varies from 0 to 1. The lower the score, the more accurate the forecasts formed by the model. Unlike tradition measures commonly used to evaluate the discrimination power of credit risk models such as ROC and CAP curve, the Brier score can be decomposed into components that suggest reasons for discrepancy and provide insights into forecast errors. This study applies the Murphy decomposition (Murphy, 1973) to decompose the Brier score. Appendix B present in details the calculation of the Brier score and its Murphy decomposition.

The first component of the Brier score, $\bar{d}(1-\bar{d})$, namely the outcome index variance, is determined by "natural forces." It reflects an aspect of forecast accuracy that does not depend on the predictive power of the estimated model (Yates, 1982, p. 139). To minimise the Brier score, it is necessary to minimise the reliability-in-the-small and to maximise the Murphy resolution.

The reliability-in-the-small, $\frac{1}{N}\sum_{j=1}^{J} N^j (f^j - \overline{d^j})^2$, measures the error that comes from the average forecast within group not measuring the average outcome within group[26]. In other words, this component measures the degree to which the $J$ distinct vector forecasts $f^j$ differs from the respective sample relative frequencies $\overline{d^j}$ ($j$=1, …, $J$) (Murphy, 1973). This term

[26] Grouping is done by sorting the holdout sample based on the survival forecast and dividing it into groups with a similar forecast. For example, observations with survival estimates varying from 0 to 10 per cent were put together, those with estimates ranging from 11 per cent to 20 per cent were in another decile and so on.

reflects a lack of model quality. The smaller the reliability-in-the-small, the lower the Brier score.

The Murphy resolution, $\frac{1}{N}\sum_{j=1}^{J} N^j(\overline{d^j}-\overline{d})^2$, measures the tendency of outcome differences in forecast groups to differ from the overall outcome. In other words, this component is a measure of the ability of the model to separate observations into $J$ sub-collections for which the sample relative frequencies $\overline{d^j}$ differ from the sample relative frequencies for the entire collection of forecasts $\overline{d}$ (Murphy, 1973). The better the information in the forecasts, the higher the Murphy resolution and the lower the Brier score.

### 7.2. Forecast horizon

The forecast horizons were chosen based on several factors. First, in practice CRAs publish rating transition matrices with a one-year horizon. Portfolio models generally use a one-year forecast horizon to calculate credit risk exposures. This horizon is also appropriate to determine regulatory capital requirements for banks (Altman, 1998). Second, the one-year horizon matches the time to events observed in the study. As indicated in Table 1, downgrade (upgrade) observations in the holdout sample mass at survival durations shorter than one year (two years). Third, previous studies suggested that the Markov property adequately holds within a one or two-year horizon (Kiefer and Larson, 2007; Frydman and Schuermann, 2008). In the light of the literature, this study focuses on evaluating the predictive accuracy of non-Markovian behaviours and time-heterogeneity within a two year forecast horizon.

The model without the current rating (Model 1) and the model with the current rating (Model 2) were employed to estimate short and intermediate term survival forecasts at different horizons within a two-year window. Due to the short term nature of the CreditWatch listing and the longer term nature of the outlook status, the model with CreditWatch (Model 4) was used to estimate short-term forecasts ($t$=0.25 year, $t$= 0.5 year) while the model with outlook (Model 3) was used to form intermediate-term forecasts ($t$=1 year, $t$=1.5 years, and $t$=2 years).

### 7.3. Forecast performance

The Brier score and its Murphy decomposition of survival estimates generated by the models for different migration outcomes $s$ at different forecast horizons $t$ are summarised in

Table 8[27]. The Brier score of each model can be assessed by reference to a naïve Brier score and a benchmark Brier score. A naïve model generating random forecasts of 0.5 has the Brier score of 0.25. The benchmark Brier score of the model for migration outcome *s* at forecast time *t* was obtained by setting the predicted survival probability $f_{s,t}$ of each holdout observation equal to the proportion of rating observations that survived from the migration outcome *s* beyond time *t* in the estimation sample.

TABLE 8 HERE

For upgrades and downgrades to high investment ratings, model 1, which includes rating history, macro-economic and political variables, exhibits good predictive accuracy. The Brier scores of Model 1 outperform the naïve and benchmark Brier scores across forecast horizons. Adding the current rating (Model 2) and the outlook (Model 3) does not change the predictive ability of the models for upgrades but improves the forecast performance of the models for downgrades to A- or higher ratings at some forecast horizons (*t*=0.25 year, *t*=1 year and *t*=1.5 years). For both migration outcomes, the model with CreditWatch (Model 4) performs poorly in comparison to the naïve model, the benchmark model, and the two estimated models (Model 1 and Model 2) as evidenced by its inferior Brier scores.

For downgrades to investment rating boundary (BBB-/BBB/BBB+), model 1 performs comparably well with the benchmark model at short term horizons but underperforms the benchmark model at intermediate term horizons. Controlling for the current rating and the outlook status improves the accuracy of survival estimates at forecast time *t*=0.5 year (model 2) and *t*=1.5 years (model 2 and model 3). Including the CreditWatch (Model 4), however, markedly reduces the predictive ability of the model, as can be seen by a substantial deterioration in the Brier scores. It is worth noting that given the rare occurrences of this downgrade outcome in the holdout sample, it is difficult to conclude on the predictive power of the estimated models.

For downgrades to speculative ratings, model 1 does not perform well compared to the benchmark model, except at *t*=0.25 years; however, it outperforms the naïve model within a 1.5 year forecast window. Adding the current rating (Model 2) raises the predictive accuracy of survival estimates as evidenced by improved (lower) Brier scores at short term horizons.

[27] To be consistent with the Murphy decomposition (Murphy, 1973) which was derived from the Sander decomposition (Sander, 1963), Table 8 reports the Sanders Brier score. This score measures the difference between a grouped survival forecast and the actual survival outcome of holdout observations (Equation A8). The difference between the reported Sanders Brier score and the Brier score calculated as in equation (A7) is minimal.

Yet, Model 2's performance deteriorates as the forecast time extends, and it substantially underperforms the naïve model at intermediate forecast horizons. Accounting for the CreditWatch listing (Model 4) yields comparable Brier scores as accounting for the current rating (Model 2).

The Murphy decompositions of the Brier scores[28] indicates that the estimated models, in most cases, offer well-calibrated survival forecasts but exhibits poor discrimination ability. This is not surprising as the migration dynamics in the estimation sample is not representative of those in the holdout sample. The financial sector was hard hit during the global financial crisis. The frequencies of downgrades to high investment ratings and downgrades to speculative ratings increased substantially during the crisis while the frequency of upgrade declined sharply (Table 7 Panel C). Benmelech and Dlugosz (2009) indicated that that rating inflation was an issue in the crisis, which certainly brings challenges to the estimated models.

The information provided by the current rating and outlook appear to be incremental compared with information contained in rating history, macro-economic factors and the political cycle. The information content of CreditWatch differs across downgrade outcomes. The forecast performance of the models for downgrades to speculative ratings suggests that CreditWatch listing seem to reflect the current decline in the credit quality of holdout observations which are heading towards this downgrade outcome. This is not surprising as these vulnerable issuers are under the scrutiny of credit rating agencies and are subject to frequent credit reviews. Rating agencies expend more resources in detecting deterioration in their credit quality and CreditWatch listing tends to be timelier when credit quality is low.

The deterioration in the predictive ability of the models for upgrades and downgrades to high/ mid investment ratings after controlling for the CreditWatch suggests that in times of market turbulence ratings are revised downward without first being placed in a negative CreditWatch. This applies to FIs, banks in particular, which have been over-rated with ratings substantially underestimating risks. Banks' high leverage and the unique nature of their assets create fundamental uncertainty for analysts (Morgan, 2002). The risks of banks' loans and trading assets are hard to observe or easy to change, and this was particularly so during the global financial crisis. Ratings tended to follow, rather than predict, the crisis (Leot, Arber, and Schou-Zibell, 2008), and rating agencies face a high reputational cost if they fail to

[28] There were only 9 observations which experienced a downgrade to investment rating boundary (BBB-/BBB/BBB+) in the holdout period and they were short-lived. No event observation survived beyond time $t=1$ year in the holdout sample. Thus, it is not possible to decompose the Brier score of survival estimates generated by the model for downgrades to investment rating boundary at intermediate- term forecast horizons.

predict imminent credit problems (Holthausen and Leftwich, 1986). If a major crisis has caught rating agencies by a surprise, they may have an incentive to be overly conservative so as to rebuild their reputation (Ferri, Liu, Stiglitz, 1999, p.352).

## 8. Conclusion

This study uses Standard & Poor's issuer rating data and develops dynamic Cox's hazard models with time-varying covariates to examine the rating migration dynamics of U.S. FIs over the period January 1984 - December 2006. The employed time-varying firm specific covariates capture the deterioration of credit quality and the effect of passing time as each FI retains its current rating grade while the time-varying macro and political cycle covariates capture the evolvement of the economic and political environment in the U.S.

The study finds that downgrade outcomes require separate models while upgrade events can be treated as equivalent within the same analysis. Different migration routes exhibit strong but markedly different within-rating dependence and time-dependence. The sources of within-rating heterogeneity can be attributed to several aspects of rating history, outlook, and CreditWatch listing whereas the sources of time-heterogeneity can be attributed to macro-economic conditions and the political cycle.

Downgrades to high investment ratings and downgrades to speculative ratings exhibit strong downward momentum. Ageing FIs or those with a high *mean rating* have a favourable experience towards upgrades, so do FIs with a history of substantial jumps to higher ratings. Issuers with a high *original rating* are at a smaller risk of falling to the investment rating threshold (BBB-/BBB/ BBB+), so do FIs with the current rating better than the average rating. FIs with a history of fallen angel events or a large rating change at lag one state are more likely to plunge to speculative ratings but less likely to descend to the investment rating boundary. FIs with a history of frequent upgrades are vulnerable to downgrades to low investment ratings and to speculative ratings. The effects of past rating behaviours discussed above persist in the presence of the current rating and the outlook/ CreditWatch listing.

Compared with some aspects of rating history such as the direction of the lagged rating change, the current rating has a relatively modest effect. However, being rated around the investment and speculative rating threshold (BBB-/BBB/BBB+, BB-/BB/BB+) strongly raises the hazard of a downgrade to speculative ratings. The *CreditWatch* status has a substantial impact in all estimated models whereas the *outlook* designation is only significant in the model for upgrades. The impact of a *negative CreditWatch* designation is particularly pronounced on downgrades to investment ratings.

In contrast to previous studies, this study finds that upgrades and downgrades to high investment ratings are more vulnerable to adverse macro-economic conditions than downgrades to low ratings. A high default rate in the financial sector reduces the probability of an upgrade but raises the hazard of a downgrade to high investment ratings. Periods characterised by an increase in *inflation expectation* and a volatile stock market saw fewer upgrades. Periods of stock market bubble observed a greater occurrence of downgrades to A- or higher while periods with a large *default spread* is associated with a lower risk of this downgrade outcome. The effects of macro-economic environment discussed above are robust after the outlook/ CreditWatch status is considered.

Upgrades and downgrades to high investment ratings are also susceptible to the political cycle. The term of a *Democratic Party's President* is associated with a lower likelihood of upgrades whereas the years when the *President's Party lost seat(s) in the mid-term congress election* is related to a higher risk of downgrades to A- or above. Controlling for the CreditWatch status, however, diminishes the impact of the political cycle.

During the holdout period (January 2007-September 2010), rating history, macro-economic and political cycle jointly exhibit good predictive accuracy for upgrades and for downgrades to high investment ratings. These factors show some calibration ability in the models for downgrades to investment boundary and downgrades to speculative ratings; however the discrimination power is rather poor. Controlling for the CreditWatch status, in most cases, does not improve the predictive power of the estimated models. Overall, the findings rule out the Markov and time-homogeneity properties inherent in the discrete time cohort Markov framework commonly used by credit rating agencies to model rating migrations.

The assessment of forecast accuracy suggests several directions in which the study may be extended. The credit quality of high rated FIs deteriorated in a dramatic manner during the global financial crisis. The information contained in the static CreditWatch/ outlook status used to form survival estimates for holdout FIs becomes increasingly stale as the survival time unfolds. One possibility to overcome this issue is to update the models using a moving window, or regularly update the time-varying covariates used to form estimates for holdout observations (Equation A5). The use of time-series macro forecasts for holdout states will control for the expected changes of macro-economic conditions over the holdout period and will introduce a forward-looking perspective into the survival estimates. Future research could also examine Standard & Poor's rating dynamics conditional on Moody's rating actions. Moody's is the rating leader of near-to-default issuers (Guttler and Wahrenburg, 2007) and its rating actions *might* trigger Standard & Poor's rating revisions for near-to-default issuers. The key results of such a study, however, are unlikely to differ from the results of this study for two reasons. First, FIs such as bank and insurance firms are inherently more opaque than other firms. Moody's and Standard & Poor's split more frequently over these financial intermediaries and the splits are more lopsided (Morgan, 2002). Second, Standard & Poor's generally assigns ratings in a timelier manner than Moody's and the tendency towards rating convergence is stronger for Moody's than for Standard & Poor's (Guttler, 2011).

The results of this study suggest that rating history, macro-economic and political cycle jointly are more important than the current rating, the outlook/ CreditWatch listing in determining future rating changes during the global financial crisis. Banks should account for these factors in assessing the credit quality of their counterparties and in determining loss reserves. The estimated dynamic hazard model can be utilised to estimate time-varying rating migration matrices for counterparties from different sectors. The dynamic model provides banks with the ability to determine dynamic economic risk capital and to detect deterioration in the credit quality of investment portfolios with a sufficient lead time. The dynamic model may also aid regulators in monitoring FIs' time-varying survival profiles and in identifying financial distressed FIs at an early stage.

## References


Allison, P. D. (1984). *Event History Analysis*, Beverly Hills, CA: Sage Publications.

Allison, P. D., Ed. (1995). *Survival Analysis Using SAS: A Practical Guide*. Cary, NC: SAS Institute Inc.

Al-Sakka, R., & Gwilym, O. (2009). Heterogeneity of sovereign rating migrations in emerging countries. *Emerging Markets Review*, 10, 151-165.

Altman, E., & Kao, D. (1991). *Corporate bond rating drift: An examination of credit quality rating changes over time.* Charlottesville, VA: The Research Foundation of the Institute of Chartered Financial Analysts.

Altman, E., & Kao, D. (1992a). Rating drift of high yield bonds. *Journal of Fixed Income,* 2, 15-20.

Altman, E., & Kao, D. (1992b). The implications of corporate bond rating drifts. *Financial Analysts Journal ,* May-June, 64-75.

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

Altman, E., & Rijken, H. (2004). How rating agencies achieve rating stability. *Journal of Banking and Finance*, 28(11), 2679-2714

Amato, J., & Furfine, C. (2004). Are Credit Ratings Procyclical?. *Journal of Banking and Finance*, 28 (11), 2641–2677

Andersen, P. K. (1992). Repeated assessment of risk factors in survival analysis. *Statistical Methods in Medical Research,*1, 297-315.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C., Schuermann, T (2002). Rating migrations and the business cycle, with applications to credit portfolio stress testing. *Journal of Banking and Finance,* 26, 445-474.

Bannier, C., & Hirsch, C. (2010). The economic function of credit rating agencies – What does the watchlist tell us? *Journal of Banking and Finance*, 34(12), 3037-3049.

Basel Committee for Banking Supervision (2005). Studies on the validation of internal rating systems. *BIS Working paper* No. 14., Bank for International Settlement, Basel.

Beck, N. (1987). Elections and the fed: Is there a political monetary cycle? *American Journal of Political Science,* 31, 194–216.

Benmelech, E. and J. Dlugosz (2009), “The Alchemy of CDO Credit Ratings,” *Journal of Monetary Economics*, 56: 615-34.

Berd, A. (2005). Dynamic estimation of credit rating transition probabilities. http://ssrn.com/abstract=842664 [accessed 30 March 2010].

Block, S., & Vaaler, P. (2004). The price of democracy: sovereign risk ratings, bond spreads and political business cycles in developing countries. *Journal of International Money and Finance*, 23, 917-946.

Blume, M., Keim, D., & Patel, S. (1991). Returns and volatility of low-grade bonds. 1977-1989. *Journal of Finance*, 46(1), 49-74.

Blume, F., Lim, F., & MacKinlay, C. (1998). The declining credit quality of US Corporate Debt: Myth or Reality? *Journal of Finance,* August(4), 1389-1413.

Bolton, P., X. Freixas and J. Shapiro (2009), “The Credit Ratings Game,” *NBER Working Paper No.14712*, Cambridge MA: National Bureau of Economic Research.

Cantor, R., & Packer, F. (1997). Differences of opinion and selection bias in the credit rating industry. *Journal of Banking and Finance*, 21, 1395-1417

Carlsen, F. (1999). Inflation and elections: theory and evidence for six OECD economies. *Economic Inquiry* 37, 120–135.

Carty, L. V., & Fons, J. S. (1994). Measuring Changes in Corporate Credit Quality. *Journal of Fixed Income*, 4(1), 27-41.

Carty, L. V. (1997). *Moody's Rating Migration and Credit Quality Correlation 1920-1996.* Special Comment (July 1997), Moody's Investor Service Global Credit Research.

Chen, L., Yen, M., Wu, H., Liao, C., Liou, D., Kuo, H., & Chen, T. H. (2005). Predictive survival model with time-dependent prognostic factors: development of computer-aided SAS Macro program. *Journal of Evaluation in Clinical Practice,* 11(2), 181-193.

Christensen, J., Hansen, E., & Lando, D. (2004). Confidence sets for continuous-time rating transition probabilities. *Journal of Banking and Finance*, 28, 2575-2602.

Cox, D. (1972). Regression models and life tables. *Journal of Royal Statistical Society* Series B (Methodological), 34, 187-220.

Dale, R., & Thomas, S. (1991). The regulatory use of credit ratings in international financial markets. *Journal of International Securities Markets*, 5(2), 9-18.

Dang, H. (2010). *Rating history, time, and the dynamic estimation of rating migration hazard*. Doctoral thesis, University of Sydney. http://hdl.handle.net/2123/6397 [accessed 30 July 2010].

Engelmann, B., Hayden, E., & Tasche, D. (2003). Measuring the discriminative power of rating systems. *Discussion paper Seties 2: Banking and Financial Supervision No. 01/2003* Deutsche Bundesbank.

Figlewski, S., Frydman, S., and Liang, W. (2012). Modeling The Effects of Macroeconomic Factors on Corporate Default and Credit Rating Transitions. *International Review of Economics and Finance*, Vol. 21, No. 1.

Fledelius, P., Lando, D., and Nielsen, J. (2004). Nonparametric Analysis of Rating Transition and Default Data. *Journal of Investment Management,* Vol. 2, No. 2.

Frydman, H., & Schuermann, T. (2008). Credit ratings dynamics and Markov mixture models. *Journal of Banking and Finance,* 32, 1062-1075.

Galil, K.( 2003). The quality of corporate credit rating: An empirical investigation. EFMA 2003 Helsinki Meetings. http://ssrn.com/abstract=406681 [accessed 30 March 2010].

Grier, K. (1989). On the existence of a political monetary cycle. *American Journal of Political Science,* 33, 376–389.

Guttler, A, & Wahrenburg, M. (2007). The adjustment of credit ratings in advance of defaults. *Journal of Banking & Finance*, 31(3), 751-767.

Guttler, A., Raupach, P. (2010). The impact of downward rating momentum. *Journal of Financial Services Research*, 37(1), 1-23.

Guttler, A. (2011). Lead-lag Relationships and Rating Convergence among Credit Rating Agencies. *Journal of Credit Risk,* Vol 7, No. 1.

Hamilton, D., & Cantor, R. (2004). *Rating transitions and default conditional on watchlist, outlook and rating history*. Special Comment Moody's Investor Service.

Haynes, S., & Stone, J. (1994). Political parties and the variable duration of business cycles. *Southern Economic Journal,* 60, 869–885.

Haynes, S., & Stone, J. (1990). Political models of the business cycle should be revived. *Economic Inquiry,* 28, 442–465.

Haynes, S., & Stone, J. (1989). An integrated test for electoral cycles in the U.S. economy. *The Review of Economics and Statistics,* 71 (3), 426–434.

Hellwig, M. (2008), "Systemic Risk in the Financial Sector: An Analysis of the Subprime-Mortgage Financial Crisis," Working Paper, Max Planck Institute for Research on Collective Goods.

Hill, P., Brooks, R., & Faff, R. (2010). Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking and Finance,* 34(6),1324-1343.

Hosmer, D., Lemeshow, S., & May, S. (2008). *Applied Survival Analysis* (2nd Ed.) New York, NY: Wiley Series in Probability and Statistics.

Hull, J. (2009), "The Credit Crunch of 2007: What Went Wrong? Why? What Lessons Can Be Learned?," *Journal of Credit Risk,* 5: 3-18.

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

Johnson, R. (2004). Rating agency actions around the investment grade boundary. *Journal of Fixed Income,* 13(4), 25-37.

Jorion, P., & Zhang, G. (2007). Information effects of bond rating changes: the role of the rating prior to the announcement. *Journal of Fixed Income*, 16, 45 – 59.

Jorion, P., Shi, C., &  Zhang, S. (2009). Tightening Credit Standards: The Role of Accounting Quality, *Review of Accounting Study*, 14(1), 123-160.

Kadam, A., & Lenk, P. (2008). Bayesian inference for issuer heterogeneity in credit rating migration.  *Journal of Banking and Finance,* 32, 2267-2274.

Kalbfleisch, J. D. and R.L. Prentice. (1980). *The Statistical Analysis Of Failure Time Data*, New York: John Wiley and Sons, Inc.

Kavvathas, D. (2001). Estimating credit rating transition probabilities for corporate bonds. AFA 2001 New Orleans Meetings. http://ssrn.com/abstract=252517 [accessed 30 March 2010].

Kiefer, N., & Larson, C. (2007). A simulation estimator for testing the time homogeneity of credit rating transitions. *Journal of Empirical Finance,* 14, 818-835.

Kim, J., & Wu, E. (2008). Sovereign credit ratings, capital flows and financial sector development in emerging markets. *Emerging Markets Review,* 9, 17-39.

Klein, M., (1996). Timing is all: elections and the duration of United States business cycles. *Journal of Money, Credit and Banking* 28 (1), 84–101.

Koopman, S., Lucas, A., & Monteiro, A .(2006). The multi-state latent factor intensity model for credit rating transitions. *Tinbergen Institute Discussion Papers* 05-071/4.

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

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

Leot, P., D. Arber and L. Schou-Zibell (2008), "Securitisation in East Asia," *Asian Development Bank Working Paper Series on Regional Economic Integration*.

Livingston, M., Naranjo, A., & Zhou, L. (2008). Split bond ratings and rating migrations. *Journal of Banking and Finance,* 32, 1613-1624.

Loeffler, G. (2005). Avoiding the rating bounce: Why rating agencies are slow to react to new information. *Journal of Economic Behaviour & Organization,* 56, 365-381.

Lown, C., Morgan, D., & Rohatgi, S. (2000). Listening to loan officers: The impact of commercial credit standards on lending and output. *Federal Reserve Bank of New York Economic Policy Review,* 6, 1-16

Lucas, D., & Lonski, J (1992). Changes in corporate credit quality 1970-1990. *Journal of Fixed Income,* 1, 7-14.

Mah, S., & Verde, M. (2004). *Rating path dependency: an analysis of corporate and structured finance rating momentum*. Fitch Ratings Report (March). New York.

Mann, C., Hamilton, D., Varma, P., & Cantor, R. (2003). *What happens to fallen angels*? A Statistical review 1982-2003. Special comment, Moody's Investor Service.

McNeil, A., & Wendin, J. (2006). Dependent Credit Migrations. *Journal of Credit Risk*, 2(3), 87-114.

Moody's Investment Service. (1994). *Note to the SEC on the disclosure of security ratings.* Moody's Investor Servives, December 5.

Morgan, D. (2002). Rating Banks: Risk and Uncertainty in an Opaque Industry. *The American Economic Review* 92(4), 874-888.

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

Pantzalis, C., Stangeland, D., & Turtle, H. (2000). Political elections and the resolution of uncertainty: the international evidence. *Journal of Banking and Finance* 24, 1575–1604.

Prentice, R., B. Williams and A. Peterson. (1981). On the regression analysis of multivariate failure time data. *Biometrika* 68(2), 373-379

Sobehart, J. R., & Keenan, S. (2001) Measuring default accurately, Credit Risk Special Report, *Risk*, 31-33.

Standard & Poor's. Dec. 12, 2001. *Playing out the credit cliff dynamic*. Standard and Poor's, New York [Re-printed 7 October 2004].

Sy, A. (2002). Emerging market bond spreads and sovereign credit ratings: reconciling market views with economic fundamentals. *Emerging Markets Review,* 3, 380-408.

Syran, R. (1991). Are we expecting a credit crunch? *New England Economic Review*, July-August, 3-10.

Tasche, D. (2008). Validation of internal rating systems and PD estimates. In G. Christodoulakis and S. Satchell (Ed.), *The Analytics of Risk Model Validation* (pp. 171-198). Elsevier.

Tsaig, Y., Levy, A., & Wang, Y. (2011). Analysing the impact of credit migration in a portfolio setting. *Journal of Banking and Finance* 35, 3145-3157.

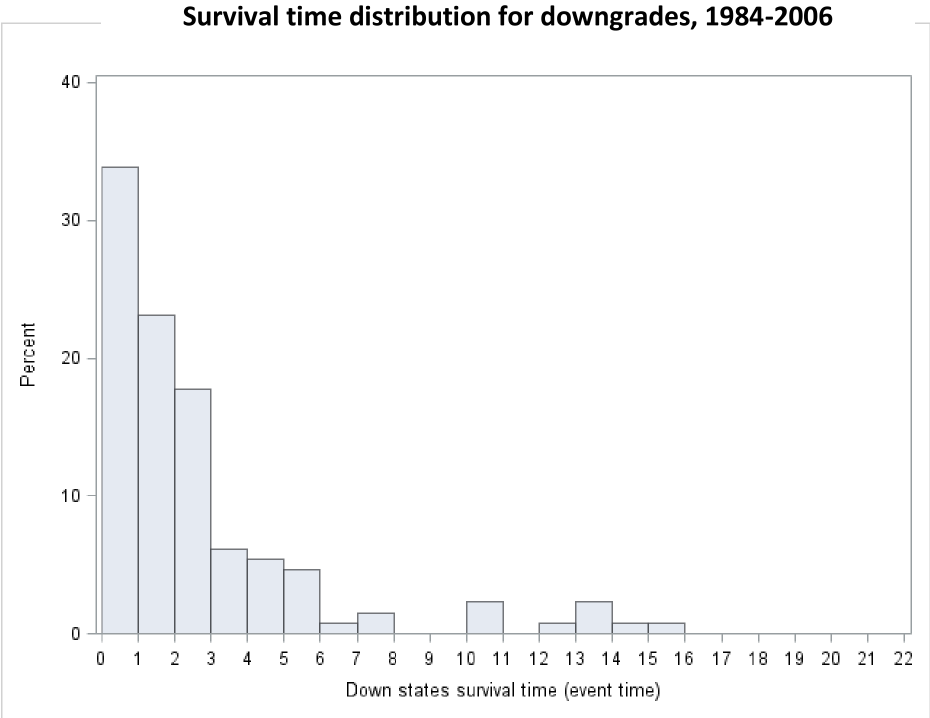
Vazza, D., Aurora, D., & Schneck, R. (2005). *Crossover Credit: A 24-Year Study of Fallen Angel Rating Behavior*, Standard & Poor's Global Fixed Income Research.

Vazza, D., Leung, E., Alsati, M., & Katz, M. (2005). *CreditWatch and ratings outlooks: Valuable predictors of rating behaviour*. Global Fixed Income Research (May). Standard & Poor's, New York .

Wei, L., Lin, D., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *Journal of the American Statistical Association,* 84, 1065-1073.

Whalen, G. (1991). A proportional hazard model of bank failure: an examination of its usefulness as an early warning tool, *Economic Review* (Q I), 21-31.

**Fig. 1**



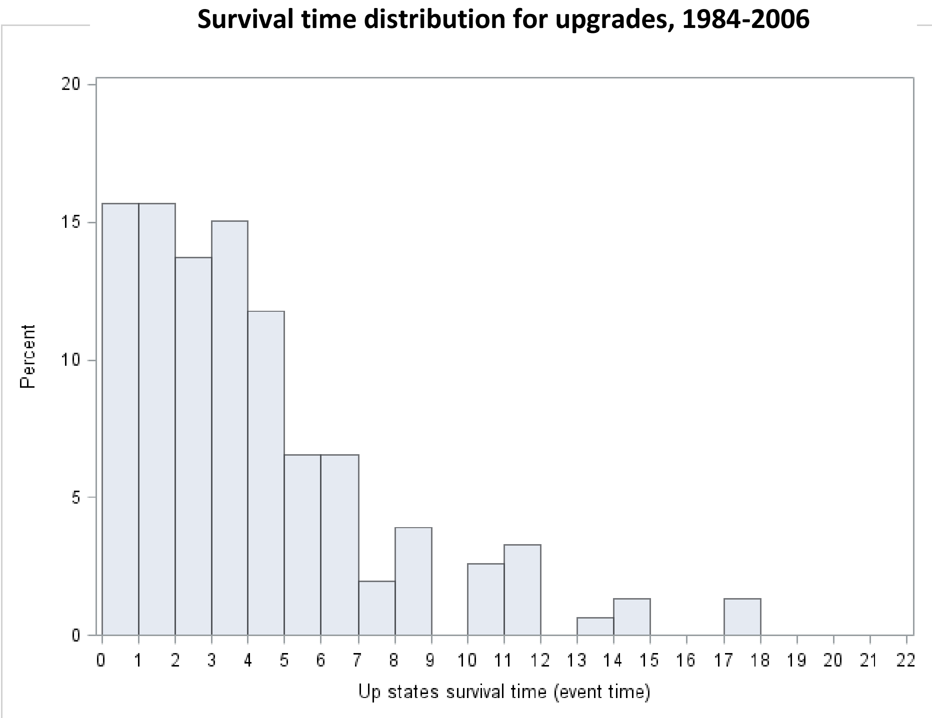




Figure 1 depicts the survival time of downgrades and upgrades in the estimation period (1984-2006). The survival time of an observation is the length of time it retains a rating grade measured from the time it enters the rating grade (*start rating)* subsequent to the commencement of the study until the time it either migrates to another rating grade (*end rating)* or becomes censored. A financial institution may contribute several rating observations to the study period. The estimation sample includes 130 downgrades and 153 upgrades

**Table 1: Descriptive statistics of survival time**

| Rating states | Sample | Number of observations | Mean (years) | Median (years) | Standard Deviation | Minimum (days) | Maximum (years) | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|
| Upgrades | Estimation | 153 | 4.18 | 3.33 | 3.45 | 22 | 17.41 | 1.525 | 2.555 |
| | Holdout | 11 | 0.50 | 0.34 | 0.52 | 10 | 1.72 | 1.4598 | 1.9917 |
| Downgrades | Estimation | 130 | 2.66 | 1.60 | 3.21 | 11 | 15.62 | 2.279 | 5.134 |
| | Holdout | 67 | 0.6028 | 0.3614 | 0.5834 | 3 | 2.8419 | 1.6905 | 2.8199 |

Table 1 above presents the descriptive statistics of the survival times for downgrades and upgrades in the estimation period (1984-2006) and in the holdout period (2007-September 2010). The survival time of an observation is the length of time it retains a rating grade measured from the time it enters the rating grade (*start rating)* subsequent to the commencement of the study until the time it either migrates to another rating grade or becomes censored. Additional analysis (not reported) indicates that down states/ up states in the estimation and the holdout periods have statistically different survival times.

## Table 2: Proportional hazard hypothesis test statistics

### Panel A: Summary of the number of rating downgrades and upgrades categorised by end ratings, 1984-2006

| | | | |
|---|---|---|---|
| Downgrades to BB- or lower ratings: | 30 | Upgrades to BB- or lower ratings : | 13 |
| Downgrades to BBB-/BBB/BBB+ : | 31 | Upgrades to BBB-/BBB/BBB+ : | 16 |
| Downgrades to A-/A/A+: | 57 | Upgrades to A-/A/A+: | 74 |
| Downgrades to AA-/AA/AA+: | 12 | Upgrades to AA-/AA/AA+/AAA : | 50 |
| Total downgrades: | 130 | Total upgrades: | 153 |

### Panel B: Maximum Likelihood Analysis of Variance

| Source | **Downgrades** Chi-Square | **Upgrades** Chi-Square |
|---|---|---|
| Intercept | 26.24*** | 5.09* |
| Survival time | 21.8*** | 7.35 |

*** p-value ≤ 1%, ** 1%< p-value ≤ 5%, * 5%< p-value ≤ 10%

### Panel C: Analysis of Maximum Likelihood Estimates

| **Parameter** | **Contrast between downgrade event types** | **Estimate** | **Contrast between upgrade event types** | **Estimate** |
|---|---|---|---|---|
| Intercept | Downgrades to BBB-/BBB/BBB+ vs Downgrades to BB+ or lower ratings | -1.3426*** | Upgrades to BB+ or lower ratings vs Upgrades to BBB-/BBB/BBB+ | 0.5749 |
| | Downgrades to A-/A/A+ vs Downgrades to BB+ or lower ratings | -1.4946*** | Upgrades to A-/A/A+ vs Upgrades to BBB-/BBB/BBB+ | 0.9847** |
| | Downgrades to AA-/AA/AA+ vs Downgrades to BB+ or lower ratings | -2.7945*** | Upgrades to AA- or higher ratings vs Upgrades to BBB-/BBB/BBB+ | 0.6929 |
| Survival time | Downgrades to BBB-/BBB/BBB+ vs Downgrades to BB+ or lower ratings | 1.2171*** | Upgrades to BB+ or lower ratings vs Upgrades to BBB-/BBB/BBB+ | -0.3085 |
| | Downgrades to A-/A/A+ vs Downgrades to BB+ or lower ratings | 1.5011*** | Upgrades to A-/A/A+ vs Upgrades to BBB-/BBB/BBB+ | 0.1409 |
| | Downgrades to AA-/AA/AA+ vs Downgrades to BB+ or lower ratings | 1.4274*** | Upgrades to AA- or higher ratings vs Upgrades to BBB-/BBB/BBB+ | 0.1186 |

*** p-value ≤ 1%, ** 1%< p-value ≤ 5%, * 5%< p-value ≤ 10%

Panel A presents the number of downgrades and upgrades categorised by end ratings. Panel B and C present the results of the multinomial logit models estimated as in Equation (1) (Cox and Oakes, 1984) for 130 downgrade events and 153 upgrade events observed in the estimation period (1984-2006). Rating observations that had not experienced a migration during the estimation period were excluded from this test against the null hypothesis that the hazards are proportional for the migration events examined. Panel B provides the analysis of variance output for the models estimated for 130 downgrades and 153 upgrades. Panel C presents the survival time coefficient estimates, derived from the maximum likelihood procedure, for (i) the contrasts between downgrades to three major investment rating groups (BBB-/BBB/BBB+, A-/A/A+, AA-/AA/AA+) and downgrade to speculative ratings (BB+ and lower ratings); (ii) the contrasts between upgrades to high investment ratings (AA-/AA/AA+/AAA), mid investment ratings (A-/A/A+) and speculative ratings (BB+ or lower ratings) versus upgrades to investment rating boundary (BBB-/BBB/BBB+). The effect of survival time $t$ for downgrades is significant at 1 per cent level (Panel B) and the beta coefficients for the contrasts between downgrade routes are statistically significant (Panel C). The log-hazards for the downgrade event contrasts diverge non-linearly with time. Both the effect of survival time $t$ (Panel B) and the beta coefficients for the contrasts between upgrade outcomes (Panel C) are not statistically significant. Thus, only the proportional hazard hypothesis for downgrades can be rejected.

## Table 3: Variable definition and references

| Variables | Definition | References |
|---|---|---|
| ***Current rating (time-independent and time-varying variables)*** | | |
| Start rating | The rating at the commencement of the current rating state | Carty and Fons (1994), Carty (1997), Hamilton and Cantor (2004), Jorion, Shi, and Zhang (2009), Figlewski et al. (2012) |
| Dummy investment boundary | The dummy takes the value of one if the start rating is in the investment grade boundary, BBB-, BBB, BBB+, and zero otherwise | Carty and Fons (1994), Carty (1997), Johnson (2004), Livingston, Naranjo, Zhou (2008) |
| Dummy junk boundary | The dummy takes the value of one if the start rating is in the speculative grade boundary, BB-, BB, BB+, and zero otherwise | |
| Dummy negative CreditWatch (time-varying variable) | The dummy takes the value of one if a firm is assigned a negative CreditWatch, and zero otherwise | |
| Dummy positive CreditWatch (time-varying variable) | The dummy takes the value of one if a firm is assigned a positive CreditWatch, and zero otherwise | Hamilton and Cantor (2004), Vazza, Leung, Alsati, and Katz (2005), Standard & Poor's (2009), Hill, Brooks, and Faff (2010), Bannier and Hirsch (2010), Guttler (2011), Al-Sakka and Gwilym (2012) |
| Dummy negative outlook (time-varying variable) | The dummy takes the value of one if a firm is assigned a negative outlook, and zero otherwise | |
| Dummy positive outlook (time-varying variable) | The dummy takes the value of one if a firm is assigned a positive outlook, and zero otherwise | |
| ***Rating history(time-independent and time-varying variables)*** | | |
| Logarithm of age since first rated (time-varying variable) | The length of time since a firm was first rated until the start of the current rating state | Altman and Kao (1991), Altman (1992), Altman (1998), Figlewski *et al*. (2012) |
| Original rating | The rating of a firm when it was first rated | Altman and Kao (1991), Altman and Kao (1992a, 1992b), Jorion *et al.* (2009), Figlewski *et al*. (2012) |
| Rating sequence | The sequence of the current rating measured since the start of the study. All observations experienced at least one rating change prior to the beginning of the study. | Dang and Partington (2008) |
| Lag one duration | The duration of the rating state that ended with either a downgrade or an upgrade and immediately preceded the current rating state | Carty and Fons (1994), Lando and Skodeberg (2002), Bannier and Hirsch (2010) |
| Lag one rating change magnitude | The magnitude of the lag one rating change, defined as lag one's end rating (i.e. the current rating) minus lag one's start rating | Guttler and Wahrenburg (2007), Al-Sakka and Gwilym (2009), Bannier and Hirsch (2010) |
| Mean rating change magnitude | The average magnitude of all rating changes occurred since the beginning of the study until the commencement of the current rating state. | |
| Dummy lag one magnitude > mean rating change magnitude | The dummy takes the value of one if the magnitude of lag one rating change is larger than the average magnitude of all rating changes , and zero otherwise | |
| Dummy lag one down | The dummy captures the direction of the lag one rating change and takes the value of one if the lag one rating ends with a downgrade, and zero otherwise | Altman and Kao (1992a,1992b), Carty and Fons (1994), Kavvathas (2001), Lando and Skodeberg (2002), Bangia Diebold, Kronimus, Schagen, and Schuermann (2002), Hamilton and Cantor (2004), Mah and Verde (2004), Figlewski *et al*. (2012) |
| Mean survival time | The average duration (survival time) a firm stayed in a rating state since the beginning of the study until the start of the current rating state | |

## Table 3: Variable definition and references (cont.)

| Variables | Definition | References |
|---|---|---|
| Dummy lag one duration > mean survival time | The dummy takes the value of one if the duration (survival time) of lag one rating is longer than the average survival time , and zero otherwise | |
| Mean rating | The average rating a firm received since the beginning of the study until the start of the current rating state | Altman and Kao (1992b), Kavvathas (2001) |
| Dummy current rating > mean rating | The dummy takes the value of one if the current rating is higher than the average rating , and zero otherwise | |
| Rate prior upgrades | This is the average number of upgrades per year over the firm's rating history. It is calculated as the number of upgrades observed between the entry of a firm to the study and the commencement of the current rating state divided by the duration from the time of entry until the start of the current rating state. | Altman and Kao (1991), Lucas and Lonski (1992), Lando and Skodeberg (2002), Koopman, Lucas, and Monteiro (2006) |
| Rate prior downgrades | This is the average number of downgrades per year over the firm's rating history. It is calculated similar to *rate prior up* except that the numerator of the ratio is the number of downgrades observed from the time the firm entered the study until the inception of the current rating state | |
| Number Fallen Angel events | The number of fallen angel events (a downgrade from an investment-grade rated rating to a speculative-grade rated rating) a firm experienced from its entry to the study until the inception of the current rating state | Mann, Hamilton, Varma, and Cantor (2003), Vazza, Aurora, and Schneck (2005) , Guttler and Wahrenburg (2007) |
| Number Rising Star events | The number of rising star events (an upgrade from a speculative-grade rated rating to an investment-grade rated rating) a firm experienced from its entry to the study until the beginning of the current rating state | |
| Number big downgrades | The number of big downgrade jumps, defined as a downgrade of at least three rating notches, a firm experienced from its entry to the study until the commencement of the current rating state | Lucas and Lonski (1992), Carty and Fons (1994) Standard and Poor's (2001), Al-Sakka and Gwilym (2009), Dang (2010) |
| Number big upgrades | The number of big upgrade jumps, defined as an upgrade of at least three rating notches, a firm experienced from its entry to the study until the inception of the current rating state | |
| ***Macro-economic (time-varying variables)*** | | |
| Dummy NBER recession | The dummy takes a value of one if an event occurs at the time of an economic recession, and zero otherwise. The economic recession is defined based on the National Bureau of Economic Research (NBER) dating of business cycle peaks (the start of recessions) and troughs (the end of recessions). | |
| Inflation expectation (%) | It is defined as the median expected price change next 12 months based on the Survey of Consumers conducted by the University of Michigan. The monthly time series were collected from the St. Louis Federal Reserve | Nickell et al. (2000), Kavvathas (2001), Bangia et al. (2002) , McNeil and Wendin (2006), Kadam and Lenk (2008), Koopman, Kraussl, Lucas, and Monteiro (2009), Bannier and Hirsch (2010), Figlewski et al. (2012) |
| Capacity utilization | This measures the extent to which the U.S. uses its productive capacity. The monthly time series were sourced from Board of Governors of the Federal Reserve System | |
| Output gap (%) | This is the deviation of the actual real GDP growth (as published quarterly by the U.S. Bureau of Economic Analysis) from the potential real GDP growth (as published quarterly by the St. Louis Federal Reserve) | |
| Standard & Poor's 500 Index return (%) | The annualised Standard & Poor's 500 Index return for a month derived from daily returns available in WRDS | |

## Table 3: Variable definition and references (cont.)

| Variables | Definition | References |
|---|---|---|
| Standard & Poor's 500 returns standard deviation (%) | Daily returns for a month are used to compute the standard deviation and this is expressed as an annual standard deviation | |
| Cyclically adjusted Price to Earnings ratio (P/E) for the aggregate stock market | This indicates the real price to earnings ratio for the overall U.S. stock market. The monthly time series were sourced from Robert Shiller's website http://www.econ.yale.edu/~shiller/data.htm | |
| Default spread (%) | This is the yield spread between Moody's Seasoned Baa Corporate Bond Yield and 10-Year Treasury Constant Maturity Rate. Both monthly time series data were collected from the St. Louis Federal Reserve | |
| Total U.S.'s corporate debt defaulting (US$ billion) | This is the total volume of corporate debt defaulting in the U.S. The yearly data was sourced from Standard and Poor's *2011 Annual U.S. Corporate Default Study And Rating Transitions* | |
| FI industry's Corporate Default rate (%) | This is the corporate default rate in the U.S. financial institutions industry. The yearly data was sourced from Standard and Poor's *2010 Annual U.S. Corporate Default Study And Rating Transitions* | |
| ***Political business cycle (time-varying variables)*** | | |
| Dummy Democratic party's President | This dummy takes a value of one if an event occurred during the term of a Democratic Party's President, and zero otherwise | www.pdamerican.org |
| Dummy presidential election year | This dummy takes a value of one if an event occurred in a year when the presidential election took place (1984, 1988, 1992, 1996, 2000, 2004, 2008), and zero otherwise | Beck (1987), Haynes et al. (1989, 1990, 1994), Klein (1996), Carlsen (1999), Pantzalis, Stangeland, and Turtle (2000), Block and Vaaler (2004) |
| Dummy President's party lost seat(s) in mid-term congress election | This dummy takes a value of one if an event occurred in a year when the President's Party lost seat(s) in the mid-term congress election, and zero otherwise | |
| Dummy President's Party not the dominant party in Congress | This dummy takes a value of one if an event occurred while the President's Party is not the dominant Party in the Congress, and zero otherwise | |

Table 3 shows the variables employed in this study. Candidate variables were screened from previous studies on credit rating migrations. Variables that exhibited strong multi-colliniarity were eliminated. Of 38 variables listed above, three time-fixed variables capturing the current rating state, four time-varying variables capture the current CreditWatch/ outlook designation, 17 time-fixed and one time-varying (*Logarithm of age since first rated*) variables capture different aspects of rating history, ten time-varying variables capture the U.S. macro-economic conditions, and four time-varying variables capture the U.S. political cycles. Time-independent variables were measured at the beginning of a rating state whereas the value of a time-varying variable used in the estimation process was updated to the most recent value as a migration event of interest occurred. The values of *Inflation expectation, Capacity utilization, Cyclically adjusted price to earnings ratio for the aggregate stock market, S&P500 Index return, S&P500 return standard deviation,* and *Default spread* were updated monthly whereas the value of *Output gap* was updated quarterly and the values of *FI Industry's corporate default rate* and *Total U.S. corporate debt defaulting* were updated yearly during the survival time of each rating observation.

## Table 4: Descriptive statistics of rating variables

| Variable | Sample | Mean | Std Dev | Minimum | Maximum | Median | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|
| Start rating (SR) | Estimation | 15.35 | 2.92 | 2 (CC) | 21 (AAA) | 16 (A) | -1.325 | 1.990 |
| | Holdout | 15.47 | 4.32 | 2 (CC) | 21 (AAA) | 17 (A+) | -1.422 | 1.528 |
| Dummy_investment_boundary | Estimation | 0.17 | 0.38 | 0 | 1 | 0 | 1.74 | 1.04 |
| (BBB-, BBB, BBB+) | Holdout | 0.1316 | 0.3395 | 0 | 1 | 0 | 2.2089 | 2.9307 |
| Dummy_junk_boundary | Estimation | 0.0649 | 0.2466 | 0 | 1 | 0 | 3.5451 | 10.6149 |
| (BB-, BB, BB+) | Holdout | 0.0702 | 0.2566 | 0 | 1 | 0 | 3.4104 | 9.8023 |
| Age_since_first_rated (years) | Estimation | 14.2860 | 8.7757 | 0.1533 | 51.7645 | 12.8624 | 1.1879 | 1.8642 |
| | Holdout | 23.3719 | 8.8061 | 1.8152 | 53.7467 | 23.4688 | 1.1133 | 3.3584 |
| Original_rating (the first rating) | Estimation | 16.2282 | 2.8945 | 8 (B+) | 21 (AAA) | 16 (A) | -0.5038 | 0.0739 |
| | Holdout | 17.0175 | 2.5486 | 9 (BB-) | 21 (AAA) | 17 (A+) | -0.4639 | -0.1667 |
| Rating sequence | Estimation | 6.0604 | 4.2037 | 2 | 25 | 5 | 1.9330 | 4.5680 |
| | Holdout | 7.2982 | 4.2840 | 2 | 25 | 6 | 1.4916 | 2.9249 |
| Lag_one_rating change magnitude | Estimation | 0.2662 | 1.7535 | -10 | 8 | 1 | 0.3907 | 3.9290 |
| | Holdout | -0.9909 | 2.0205 | -8 | 8 | -1 | 1.3059 | 7.4654 |
| Mean rating change magnitude | Estimation | 0.0023 | 1.1041 | -3.0315 | 5.0000 | -0.0480 | 0.8040 | 2.1000 |
| | Holdout | -0.1402 | 1.0092 | -3.0000 | 2.2658 | -0.0082 | -0.5957 | 0.0829 |
| Dummy_lag one magnitude > mean magnitude | Estimation | 0.4407 | 0.4970 | 0 | 1 | 0 | 0.2396 | -1.9513 |
| | Holdout | 0.2000 | 0.4018 | 0 | 1 | 0 | 1.5208 | 0.3183 |
| Lag one duration (years) | Estimation | 4.0363 | 3.8852 | 0.0301 | 27.6797 | 2.8172 | 1.9580 | 5.1440 |
| | Holdout | 3.6379 | 5.7968 | 0.0082 | 23.4552 | 1.0705 | 2.0156 | 2.9100 |
| Mean prior rating duration | Estimation | 3.9570 | 3.2721 | 0.1533 | 27.6797 | 2.8000 | 2.7230 | 11.0080 |
| | Holdout | 4.9944 | 3.2895 | 1.5746 | 21.6783 | 3.9115 | 2.3298 | 6.9256 |
| Dummy_lag one duration > mean prior rating duration | Estimation | 0.3937 | 0.4891 | 0 | 1 | 0 | 0.4365 | -1.8176 |
| | Holdout | 0.2018 | 0.4031 | 0 | 1 | 0 | 1.5063 | 0.2733 |
| Mean rating | Estimation | 15.6429 | 2.4685 | 7.2448 | 21 (AAA) | 16 (A) | -0.6427 | 0.3771 |
| | Holdout | 16.6793 | 1.8063 | 9.7205 | 21 (AAA) | 16.8339 | -1.0345 | 2.9825 |
| Dummy current rating > mean rating | Estimation | 0.4743 | 0.4999 | 0 | 1 | 0 | 0.1034 | -1.9983 |
| | Holdout | 0.4825 | 0.5019 | 0 | 1 | 0 | 0.0712 | -2.0309 |
| Dummy_lag one_down | Estimation | 0.4474 | 0.4978 | 0 | 1 | 0 | 0.2122 | -1.9638 |
| | Holdout | 0.7807 | 0.4156 | 0 | 1 | 1 | -1.3749 | -0.1117 |
| Number_Fallen angel events | Estimation | 0.2170 | 0.5356 | 0 | 3 | 0 | 2.5963 | 6.4257 |
| | Holdout | 0.3070 | 0.5174 | 0 | 2 | 0 | 1.4197 | 1.0926 |
| Number_Rising star events | Estimation | 0.1633 | 0.4568 | 0 | 3 | 0 | 3.0108 | 9.2169 |
| | Holdout | 0.1754 | 0.3820 | 0 | 1 | 0 | 1.7295 | 1.0086 |
| Number_big_downgrades | Estimation | 0.0268 | 0.1618 | 0 | 1 | 0 | 5.8744 | 32.6550 |
| | Holdout | 0.1491 | 0.3578 | 0 | 1 | 0 | 1.9964 | 2.0208 |
| Number_big_upgrades | Estimation | 0.2125 | 0.5619 | 0 | 3 | 0 | 2.7540 | 7.0456 |
| | Holdout | 0.1228 | 0.3555 | 0 | 2 | 0 | 2.9127 | 8.3392 |
| Rate_prior_upgrades | Estimation | 0.1735 | 0.3337 | 0 | 6.5223 | 0.1402 | 15.7250 | 295.6699 |
| | Holdout | 0.1156 | 0.0803 | 0 | 0.2729 | 0.1017 | 0.0987 | -1.1670 |
| Rate_prior_downgrades | Estimation | 0.2295 | 0.1951 | 0 | 1.2998 | 0.1945 | 1.8041 | 5.5340 |
| | Holdout | 0.1467 | 0.0869 | 0 | 0.5509 | 0.1353 | 1.1697 | 3.3141 |

Table 4 reports the descriptive statistics of rating variables for 447 observations in the estimation period (1984-2006) and 114 observations in the holdout period (2007-September 2010). *Start rating* is the current rating grade. *Dummy investment boundary*/ *Dummy junk boundary* indicates whether a firm is BBB-,BBB, BBB+ rated/ BB-, BB, BB+ rated. *Age since first rated* is the number of years spanning from the time a firm was first rated till the beginning of the current rating state. *Original rating* is the rating received when a firm was first rated. *Rating sequence* is the sequence of the current rating state. *Lag one rating change magnitude* is the magnitude of the rating change preceding the current rating. *Mean rating change magnitude* is the average of the magnitudes of the rating changes preceding the current rating. *Dummy lag one magnitude > mean magnitude* indicates whether the magnitude of lag one change is larger than the average magnitude. *Lag one duration* is the duration of the rating change that precedes the current rating. *Mean prior rating duration* is the average survival time a firm stayed in a rating state. *Dummy lag one duration > mean prior rating duration* indicates whether the survival duration of lag one rating is longer than the average survival duration of prior ratings. *Mean rating* is the average rating a FI has received. *Dummy current rating > mean rating* indicates if the current rating grade is better than the average rating. *Dummy lag one down* indicates if the lag one rating ends with a downgrade. *Number fallen angel events*/ *Number rising star events* is the number of times a firm experienced a downgrade from investment ratings to junk ratings/ an upgrade from junk ratings to investment ratings. *Number big downgrades*/ *Number big upgrades* is the number of times a firm experienced a substantial downgrade/ upgrade of at least three rating notches. *Rate prior upgrades*/ *Rate prior downgrades* is the average number of upgrades/ downgrades a firm experienced in a year prior to the current rating.

## Table 5: Descriptive statistics of CreditWatch (CW) and Outlook (OL), 1984-2006

### Panel A: CreditWatch duration and history

| | | Duration of CWs (years) | | | | Number of prior negative CWs | | | | Number of prior positive CWs | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total obs | Mean | Std Dev | Minimum | Maximum | Mean | Std Dev | Minimum | Maximum | Mean | Std Dev | Minimum | Maximum |
| All rating-CreditWatch observations | 686 | 1.455 | 1.931 | 2 days | 13.541 | 0.375 | 0.588 | 0 | 4 | 0.292 | 0.461 | 0 | 2 |
| All rating- CreditWatch observations with positive CreditWatch(s) | 166 | 0.676 | 0.734 | 9 days | 4.966 | 0.048 | 0.215 | 0 | 1 | 1.012 | 0.109 | 1 | 2 |
| All rating- CreditWatch observations with negative CreditWatch(s) | 161 | 0.690 | 0.854 | 6 days | 4.107 | 1.168 | 0.451 | 1 | 4 | 0.019 | 0.136 | 0 | 1 |

### Panel B: Outlook duration and history

| | | Duration of Ols (years) | | | | Number of prior negative OLs | | | | Number of prior positive OLs | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total obs | Mean | Std Dev | Minimum | Maximum | Mean | Std Dev | Minimum | Maximum | Mean | Std Dev | Minimum | Maximum |
| All rating-outlook observations | 1157 | 1.439 | 1.686 | 2 days | 12.676 | 0.509 | 0.788 | 0 | 5 | 0.275 | 0.501 | 0 | 2 |
| All ratings-outlook observations with positive outlook(s) | 135 | 0.985 | 0.871 | 2 days | 5.218 | 0.244 | 0.717 | 0 | 5 | 1.104 | 0.306 | 1 | 2 |
| All ratings-outlook observations with negative outlook(s) | 209 | 1.066 | 0.862 | 3 days | 3.775 | 1.325 | 0.679 | 1 | 5 | 0.148 | 0.407 | 0 | 2 |

Table 5 reports the descriptive statistics of the duration and history of CreditWatch/ Outlook designations for rating-CreditWatch/ rating-outlook observations in the estimation period (1984-2006). CreditWatch designations may be Watch Developing, Watch Negative, and Watch Positive. Outlook designations may be Developing, Stable, Negative and Positive. An issuer placed on CreditWatch does not carry an outlook during the CreditWatch review period. CreditWatch status generally lasts for up to 90 days whereas outlook indicates the potential direction of a long-term credit rating over the intermediate term (typically six months to two years). Outlook generally covers up to two years for investment grade and one year for junk grade. A positive (negative) CreditWatch/ outlook designation means that a rating may be raised (lowered) while the developing CreditWatch/ outlook designation means that a rating may be raised, lower or affirmed. A stable outlook means that a rating is not likely to change. An outlook is not necessarily a precursor of a rating change or future CreditWatch action, and does not mean that an issuer has unfavourable credit characteristics. A CreditWatch listing does not mean a rating change is inevitable, and a rating change can occur without the rating being placed on CreditWatch beforehand.

## Table 6: Descriptive statistics of macro-economic variables

| Variables | Sample | Mean | Median | Std Dev | Minimum | Maximum | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|
| Capacity utilization | Estimation | 80.690 | 80.823 | 2.717 | 74.856 | 84.609 | -0.533 | -0.615 |
| | Holdout | 76.084 | 78.393 | 4.559 | 69.451 | 80.402 | -0.394 | -1.702 |
| Inflation expectation (%) | Estimation | 3.052 | 2.979 | 0.414 | 2.172 | 4.203 | 0.514 | 0.068 |
| | Holdout | 3.256 | 3.193 | 0.468 | 2.665 | 4.297 | 0.905 | 0.036 |
| Output gap (%) | Estimation | -44.042 | -58.291 | 130.400 | -348.495 | 276.540 | 0.453 | -0.181 |
| | Holdout | -375.67 | -202.77 | 419.020 | -957.787 | 75.703 | -0.326 | -1.705 |
| Cyclically adjusted price to earnings ratio for the aggregate stock market | Estimation | 22.914 | 21.014 | 8.923 | 9.415 | 42.743 | 0.536 | -0.438 |
| | Holdout | 22.181 | 22.715 | 3.921 | 16.774 | 26.816 | -0.083 | -1.780 |
| S&P500 Index return (%) | Estimation | 0.829 | 0.889 | 1.138 | -2.909 | 3.060 | -0.074 | 0.652 |
| | Holdout | -0.307 | 0.4867 | 1.949 | -4.865 | 2.214 | -1.1909 | 0.3967 |
| S&P500 return standard deviation (%) | Estimation | 3.063 | 2.769 | 1.012 | 1.638 | 5.788 | 0.690 | -0.538 |
| | Holdout | 4.939 | 4.370 | 2.216 | 1.973 | 8.960 | 0.466 | -0.991 |
| Default spread (%) | Estimation | 0.950 | 0.888 | 0.260 | 0.592 | 1.518 | 0.489 | -0.973 |
| | Holdout | 1.456 | 1.309 | 0.530 | 0.888 | 2.477 | 0.612 | -1.005 |
| Total U.S. corporate debt defaulting (US$ billion) | Estimation | 24.51 | 7.28 | 41.71 | 0.31 | 188.14 | 2.852 | 8.084 |
| | Holdout | 188.83 | 7.02 | 210.45 | 6.97 | 516.08 | 0.4523 | -1.5393 |
| FI Industry's corporate default rate (%) | Estimation | 0.77 | 0.31 | 0.96 | 0.00 | 2.75 | 1.0814 | -0.4037 |
| | Holdout | 1.69 | 0.6 | 1.60 | 0.00 | 3.78 | 0.2655 | -1.7726 |

Table 6 reports the descriptive statistics of the time series for the exponentially weighted averages of the macro-economic variables in the estimation period (1984-2006) and in the holdout period (2007-September 2010). Additional analysis (not reported) show that the values of macro-economic variables in the two periods are statistically different. *Capacity utilization* measures the extent to which the U.S. uses its productive capacity. *Inflation expectation* is the median expected price change next 12 months based on the Survey of Consumers conducted by the University of Michigan. *Output gap* measures the deviation of the actual real GDP growth from the potential real GDP growth. *Cyclically adjusted price to earnings ratio for the aggregate stock market* indicates the real price to earnings ratio for the overall U.S. stock market. *S&P500 Index return* is the annualised Standard & Poor's 500 Index return for a month derived from daily returns. *S&P500 return standard deviation* is the annualised standard deviation of the Standard & Poor's 500 Index return derived from daily returns in each month. *Default spread* is the yield spread between Moody's Seasoned Baa Corporate Bond Yield and 10-Year Treasury Constant Maturity Rate. *FI Industry's corporate default rate* is the corporate default rate in the U.S. financial institutions industry. *Total U.S. corporate debt defaulting* is the volume (US$ billion) of corporate debt default in the U.S. Except *Dummy NBER recession* (not reported in the above table), macro-economic variables were constructed as exponentially weighted averages of lagged observations computed monthly over an 18-month window. The construction of lagged values is similar to Figlewski *et al.* (2012)'s approach.

## Table 7: Estimation models, 1984-2006

### Panel A: Parameter estimates

| | Upgrades - All end ratings | | | | Downgrades categorised by end ratings: Downgrades to A- or higher rating grades | | | | Downgrades to BBB-, BBB, BBB+ | | | | Downgrades to junk ratings | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, CW (4) |
| Variables | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate |
| ***Current rating state*** | | | | | | | | | | | | | | | |
| Start rating (SR) | NA | -0.3777*** | -0.3251*** | -0.31262*** | NA | 1.1191*** | 1.1185*** | 0.8635*** | NA | 0.3245*** | 0.2962*** | 0.4974*** | NA | -0.6391*** | -0.6879*** |
| Dummy investment boundary | NA | | | | NA | NA | NA | NA | NA | 1.124*** | 0.8194* | 2.1151*** | NA | 3.3283** | 4.0422** |
| Dummy junk boundary | NA | | | | NA | NA | NA | NA | NA | NA | NA | NA | NA | 3.4448*** | 4.2139*** |
| Dummy negative Outlook | NA | NA | | NA | NA | NA | | NA | NA | NA | | NA | NA | NA | NA |
| Dummy positive Outlook | NA | NA | 1.1293*** | NA | NA | NA | | NA | NA | NA | | NA | NA | NA | NA |
| Dummy negative CreditWatch | NA | NA | NA | | NA | NA | NA | 3.7485*** | NA | NA | NA | 5.3621*** | NA | NA | 1.8759*** |
| Dummy positive CreditWatch | NA | NA | NA | 3.3193*** | NA | NA | NA | | NA | NA | NA | | NA | NA | |
| ***Rating history*** | | | | | | | | | | | | | | | |
| Rating sequence | | | | | 0.1261* | 0.256*** | 0.2467*** | | 0.1612** | | 0.1405* | | | | |
| Log age | 1.0303*** | 0.6969*** | 0.5816*** | 0.6786*** | | | | | 1.1633** | 1.0968** | 1.0278** | | | | |
| Original rating | | | | -0.1087* | -0.4738*** | | | | -0.3041*** | -0.3565*** | -0.45*** | -0.6399*** | -0.309** | | |
| Lag one rating change magnitude | -0.2509*** | | | | 0.3283*** | | | | | -0.5548*** | -0.5833*** | -0.8001*** | 0.4558*** | 0.5256*** | 0.5679** |
| Lag one duration | | | | | | | | | | | | | | | 0.5249** |
| Dummy lag one down | -1.1103*** | -0.9227*** | -0.8391*** | -0.6834*** | 0.8924*** | 1.7111*** | 1.6774*** | 1.6723*** | 2.1091*** | | | | 3.3125*** | 4.1368*** | 6.6658*** |
| Mean rating change magnitude | -0.3957*** | | | -0.2674* | 0.6494*** | | | | -0.8991** | -0.5579** | -0.859** | | -0.7741** | | |
| Dummy lag one magnitude > mean rating change magnitude | | | | | | | | | | | | | | | 2.3275** |
| Mean survival time | | | | | | | | | | | | | | | -0.4759** |
| Dummy lag one duration > mean survival time | | | | | | | | | | | | | | | -2.3371* |
| Mean rating | -0.1578*** | 0.1826*** | 0.1687*** | 0.2620*** | 1.2972*** | | | -0.2492** | | | | | | | |
| Dummy start rating > mean rating | NA | | | | NA | | | | NA | -1.4839*** | -1.3164** | -2.5173*** | NA | | |
| Number Fallen Angel events | | -0.5543** | | | | | | | -1.4608*** | | -1.1992** | -2.5909*** | 3.1537*** | 0.7993*** | 1.223*** |
| Number Rising Star events | | 0.5836** | | | 2.2418*** | | | | | | | 3.0221*** | -2.3122** | | |
| Number big downgrade events | | | | | | | | | | | | | | | |
| Number big upgrade events | | 0.4455*** | 0.5065*** | 0.3768** | | | | | | | | | -0.7433** | | |

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

**Table 7: Estimation models, 1984-2006**

**Panel A: Parameter estimates (cont.)**

| | **Upgrades - All end ratings** | | | | **Downgrades to A- or higher rating grades** | | | | **Downgrades to BBB-, BBB, BBB+** | | | | **Downgrades to junk ratings** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, CW (4) |
| Variables | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate | Parameter estimate |
| Rate prior upgrades | | | | | | | | | 1.4034*** | 1.3387*** | 1.3994*** | 0.5696*** | | 1.2461*** | 1.564*** |
| Rate prior downgrades | 2.8093*** | 2.377*** | 2.2012*** | | | -1.9591* | -1.9911* | | | | | | | | |
| ***Macro-economic*** | | | | | | | | | | | | | | | |
| FI sector's default (%) | -0.4383** | -0.5015*** | -0.414** | -0.6814*** | 1.5729*** | 1.741*** | 1.6526*** | 0.9037*** | 0.6586* | | | | -1.0127*** | | |
| $US billion corporate debt default | | | | | | | | | 0.0166*** | 0.0111*** | 0.0076** | 0.0151*** | | | |
| Default spread (%) | 4.1531*** | 4.5498*** | 4.6042*** | | -4.1941*** | -6.0943*** | -5.3581*** | -5.8975*** | | | | | | | |
| Capacity utilization | 0.3238*** | 0.3423*** | 0.3637*** | | | | | -0.435*** | | | | | | | |
| Output gap (%) | | | | 0.002* | | | | | | | | 0.0041*** | | | |
| Dummy NBER recession | 1.4595** | 1.4762** | 1.6069** | | | 1.0244** | | | | | | | | | |
| Inflation expectation (%) | -2.1763*** | -2.2791*** | -2.3972*** | -1.526*** | | | | | | | | -1.4947* | 1.382** | 1.172** | |
| SP500 Index return (%) | | | | | -0.3747*** | -0.4109*** | -0.3996*** | | | | | | -0.5591*** | -0.5258*** | -0.5182*** |
| SP500 standard deviation | -1.0109*** | -1.0583*** | -1.0302*** | -0.4718*** | | | | | | | | | | | |
| P/E ratio for the aggregate stock market | 0.0848*** | 0.1042*** | 0.1073*** | | 0.196*** | 0.1681*** | 0.1758*** | 0.158*** | | | | | | 0.0901*** | 0.1387*** |
| ***Political cycle*** | | | | | | | | | | | | | | | |
| Dummy presidential election year | | | | | | -0.7646* | | -1.5681*** | 0.8298* | 0.8036* | | | | | |
| Dummy Democratic party's President | -1.3198** | -1.413*** | -1.3654** | -1.1475*** | | | | | 1.5172* | | | | | | |
| Dummy President's party lost seat in mid-term Congress election | | | | | 3.2539*** | 3.2551*** | 3.5217*** | 1.4111** | 1.3508* | 1.025 | | | | | 1.2447 |
| Dummy President's party not the dominant party in the Congress | -0.5326** | -0.5482** | -0.6154*** | | -2.0104*** | -2.0182*** | -2.045*** | | | | | | | | |

Panel A reports the beta coefficients of the significant variables in the generic upgrade models and the downgrade models for downgrades to A- or higher ratings (stratified by end ratings), downgrades to investment rating boundary BBB-/ BBB/ BBB+, and downgrades to junk rating grades. The generic upgrade models treat all upgrades as equivalent events. Each type specific downgrade hazard model treats the downgrades being modelled as events, and treats other downgrade events and survival states as censored. The model without start rating (model 1) includes 17 rating history variables, 10 time-varying macro-economic variables, and four time-varying political cycle variables. The model with start rating (model 2) includes start rating variable(s), 17 rating history variables as in model 1 plus *Dummy start rating > mean rating*, and 14 time-varying macro-economic and political cycle variables as in model 1. The model with start rating and outlook (model 3) includes all variables in model 2 and two time-varying outlook dummy variables. The model with start rating and CreditWatch (model 4) includes all variables in model 2 and two time-varying CreditWatch dummy variables. For downgrades to junk ratings, only the results of models 1, 2 and 4 were presented as model 3 (with SR and OL) was the same as model 2 (with SR). The backward selection procedure was employed. Variables were retained the models according to the log-likelihood ratio test, at the 10 per cent level or better, derived from the maximum likelihood procedure used to estimate the models. Parameter estimates are given first followed by the corresponding p-values based on Wald chi-square tests (*** p-value ≤ 1%, ** 1%< p-value ≤ 5%, * 5%< p-value ≤ 10%).

**Table 7: Estimation models, 1984-2006**

**Panel B: Model goodness of fit, 1984-2006**

| | Upgrades - All end ratings | | | | Downgrades categorised by end ratings: Downgrades to A- or higher ratings | | | | Downgrades to BBB-, BBB, BBB+ | | | | Downgrades to junk ratings | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model w/o SR (1) | Model w SR (2) | Model w SR,OL (3) | Model w SR,CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR,CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR,CW (4) | Model w/o SR (1) | Model w SR (2) | Model w SR, CW (4) |
| -2 Log likelihood (without covariates) | 1522.66 | 1522.66 | 1522.66 | 1522.66 | 579.498 | 579.498 | 579.498 | 579.498 | 343.603 | 343.60 | 343.603 | 343.603 | 353.620 | 353.620 | 353.620 |
| -2 Log likelihood (with covariates) | 1386.58 | 1363.27 | 1339.69 | 1095.42 | 429.802 | 405.55 | 410.176 | 307.099 | 273.104 | 274.98 | 274.851 | 165.289 | 202.836 | 185.713 | 166.137 |
| Likelihood ratio Chi-square | 136.081 | 159.39 | 182.978 | 427.243 | 149.7 | 173.95 | 169.32 | 272.4 | 70.499 | 68.63 | 68.752 | 177.774 | 150.78 | 167.906 | 187.483 |
| Degrees of freedom | 15 | 17 | 16 | 13 | 13 | 12 | 10 | 10 | 12 | 11 | 11 | 12 | 10 | 10 | 15 |
| Pr > ChiSq | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |

Panel B reports the model fit statistics. The term *-2 Log-Likelihood* is the logarithm of the maximum likelihood estimator for the estimated model. The likelihood ratio (*LR*) is calculated as $LR=2(\ln L_1 - \ln L_o)$ where $L_1$ is the log-likelihood of the estimated model and $L_o$ is the log-likelihood of the model without covariates. The likelihood ratio test has an asymptotic chi-square distribution where the degree of freedom is the number of additional parameters in the estimated model. Comparison of the log-likelihood statistics in Panel B shows that the explanatory power of each model is significantly improved as variables are added. The likelihood ratio reports that this improvement is significant at better than the 1 per cent level.

**Panel C: Summary of the Number of Event and Censored Observations**

| | Upgrades - All end ratings | | | | Downgrades categorised by end ratings: Downgrades to A- or higher rating grades | | | | Downgrades to BBB-, BBB, BBB+ | | | Downgrades to junk rating grades | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | Event | Censored | % event | Total | Event | Censored | % event | Event | Censored | % event | Event | Censored | % event |
| Estimation sample | 447 | 153 | 294 | 34.23% | 447 | 69 | 378 | 15.44% | 31 | 416 | 6.94% | 30 | 417 | 6.71% |
| Holdout sample | 114 | 11 | 103 | 9.65% | 114 | 39 | 75 | 34.21% | 9 | 105 | 7.89% | 19 | 95 | 16.67% |

Panel C reports the number of rating observations, the number of events and censored observations for each event type examined in the estimation period (1984-2006) and in the holdout period (2007-September 2010). Censored observations in each model consist of incomplete durations, rating withdrawals, and all migration types other than those being modelled in that model. Additional analysis (not reported) shows that the proportion of migration events examined in the estimation period (1984-2006) and in the holdout period (2007-September 2010) are statistically different.

**Table 8: Brier scores for the probability survival estimates of holdout rating observations, January 2007-September 2010**

**Upgrades-All end ratings combined**

| Short-term forecasts | | Benchmark (Empirical) | Naïve | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) |
|---|---|---|---|---|---|---|---|
| t=0.25 year | **Brier score** | 0.1476 | 0.25 | **0.0965** | **0.0965** | NA | **0.5726** |
| (114 obs) | Outcome index variance | | | 0.0872 | 0.0872 | NA | 0.0872 |
| | Murphy resolution | | | 0.0000 | 0.0000 | NA | 0.0026 |
| | Reliability-in-the-small | | | 0.0093 | 0.0093 | NA | 0.488 |
| t=0.5 year | **Brier score** | 0.1405 | 0.25 | **0.0740** | **0.0736** | NA | **0.9259** |
| (81 obs) | Outcome index variance | | | 0.0686 | 0.0686 | NA | 0.0686 |
| | Murphy resolution | | | 0.0000 | 0.0000 | NA | 0.0000 |
| | Reliability-in-the-small | | | 0.0054 | 0.005 | NA | 0.8573 |
| Intermediate-term forecasts | | | | | | | |
| t=1 year | **Brier score** | 0.141 | 0.25 | **0.0724** | **0.0737** | **0.0739** | NA |
| (53 obs) | Outcome index variance | | | 0.0698 | 0.0698 | 0.0698 | NA |
| | Murphy resolution | | | 0.0011 | 0.0011 | 0.0011 | NA |
| | Reliability-in-the-small | | | 0.0037 | 0.005 | 0.0052 | NA |
| t=1.5 years | **Brier score** | 0.1338 | 0.25 | **0.053** | **0.0527** | **0.0528** | NA |
| (38 obs) | Outcome index variance | | | 0.0499 | 0.0499 | 0.0499 | NA |
| | Murphy resolution | | | 0.0003 | 0.0003 | 0.0004 | NA |
| | Reliability-in-the-small | | | 0.0034 | 0.0031 | 0.0033 | NA |
| t=2 years | **Brier score** | 0.1357 | 0.25 | **0.0595** | **0.0589** | **0.0589** | NA |
| (17 obs) | Outcome index variance | | | 0.0554 | 0.0554 | 0.0554 | NA |
| | Murphy resolution | | | 0.0011 | 0.0007 | 0.0007 | NA |
| | Reliability-in-the-small | | | 0.0052 | 0.0042 | 0.0042 | NA |

**Downgrades to A- or higher rating grades**

| Short-term forecasts | | Benchmark (Empirical) | Naïve | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) |
|---|---|---|---|---|---|---|---|
| t=0.25 year | **Brier score** | 0.2603 | 0.25 | **0.1752** | **0.1515** | NA | **0.8246** |
| (114 obs) | Outcome index variance | | | 0.1447 | 0.1447 | NA | 0.1447 |
| | Murphy resolution | | | 0.0000 | 0.0284 | NA | 0.0000 |
| | Reliability-in-the-small | | | 0.0305 | 0.0352 | NA | 0.6799 |
| t=0.5 year | **Brier score** | 0.2799 | 0.25 | **0.1345** | **0.2196** | NA | **0.8642** |
| (81 obs) | Outcome index variance | | | 0.1174 | 0.1174 | NA | 0.1174 |
| | Murphy resolution | | | 0.0001 | 0.0022 | NA | 0.0000 |
| | Reliability-in-the-small | | | 0.0172 | 0.1044 | NA | 0.7468 |
| Intermediate-term forecasts | | | | | | | |
| t=1 year | **Brier score** | 0.2716 | 0.25 | **0.1810** | **0.1573** | **0.1543** | NA |
| (53 obs) | Outcome index variance | | | 0.1531 | 0.1531 | 0.1531 | NA |
| | Murphy resolution | | | 0.0168 | 0.0126 | 0.0168 | NA |
| | Reliability-in-the-small | | | 0.0447 | 0.0168 | 0.0180 | NA |
| t=1.5 years | **Brier score** | 0.2057 | 0.25 | **0.1315** | **0.1151** | **0.1138** | NA |
| (38 obs) | Outcome index variance | | | 0.1143 | 0.1143 | 0.1143 | NA |
| | Murphy resolution | | | 0.0000 | 0.0332 | 0.0332 | NA |
| | Reliability-in-the-small | | | 0.0172 | 0.0340 | 0.0327 | NA |
| t=2 years | **Brier score** | 0.2271 | 0.25 | **0.0865** | **0.2942** | **0.2940** | NA |
| (17 obs) | Outcome index variance | | | 0.0554 | 0.0554 | 0.0554 | NA |
| | Murphy resolution | | | 0.0011 | 0.0019 | 0.0011 | NA |
| | Reliability-in-the-small | | | 0.0322 | 0.2407 | 0.2397 | NA |

**Downgrades to speculative rating grades**

| Short-term forecasts | | Benchmark (Empirical) | Naïve | Model w/o SR (1) | Model w SR (2) | Model w SR, CW (4) |
|---|---|---|---|---|---|---|
| t=0.25 year | **Brier score** | 0.1488 | 0.25 | **0.1491** | **0.1324** | **0.1339** |
| (114 obs) | Outcome index variance | | | 0.1389 | 0.1389 | 0.1389 |
| | Murphy resolution | | | 0.0134 | 0.0204 | 0.0178 |
| | Reliability-in-the-small | | | 0.0236 | 0.0139 | 0.0128 |
| t=0.5 year | **Brier score** | 0.1114 | 0.25 | **0.1817** | **0.1019** | **0.1045** |
| (81 obs) | Outcome index variance | | | 0.1082 | 0.1082 | 0.1082 |
| | Murphy resolution | | | 0.0148 | 0.0204 | 0.0248 |
| | Reliability-in-the-small | | | 0.0883 | 0.0141 | 0.0211 |
| Intermediate-term forecasts | | | | | | |
| t=1 year | **Brier score** | 0.0698 | 0.25 | **0.1400** | **0.3153** | NA |
| (53 obs) | Outcome index variance | | | 0.0698 | 0.0698 | NA |
| | Murphy resolution | | | 0.0088 | 0.0015 | NA |
| | Reliability-in-the-small | | | 0.079 | 0.247 | NA |
| t=1.5 years | **Brier score** | 0.0501 | 0.25 | **0.1584** | **0.4525** | NA |
| (38 obs) | Outcome index variance | | | 0.0499 | 0.0499 | NA |
| | Murphy resolution | | | 0.005 | 0.0051 | NA |
| | Reliability-in-the-small | | | 0.1135 | 0.4077 | NA |
| t=2 years | **Brier score** | 0.0554 | 0.25 | **0.3326** | **0.7783** | NA |
| (17 obs) | Outcome index variance | | | 0.0554 | 0.0554 | NA |
| | Murphy resolution | | | 0.0024 | 0.0063 | NA |
| | Reliability-in-the-small | | | 0.2796 | 0.7292 | NA |

**Downgrades to BBB-, BBB, BBB+**

| Short-term forecasts | | Benchmark (Empirical) | Naïve | Model w/o SR (1) | Model w SR (2) | Model w SR, OL (3) | Model w SR, CW (4) |
|---|---|---|---|---|---|---|---|
| t=0.25 year | **Brier score** | 0.0728 | 0.25 | **0.0734** | **0.0924** | NA | **0.8842** |
| (114 obs) | Outcome index variance | | | 0.0727 | 0.0727 | NA | 0.0727 |
| | Murphy resolution | | | 0.0000 | 0.0000 | NA | 0.0003 |
| | Reliability-in-the-small | | | 0.0007 | 0.0197 | NA | 0.8118 |
| t=0.5 year | **Brier score** | 0.0686 | 0.25 | **0.0701** | **0.0610** | NA | **0.9259** |
| (81 obs) | Outcome index variance | | | 0.0686 | 0.0686 | NA | 0.0686 |
| | Murphy resolution | | | 0.0000 | 0.0094 | NA | 0.0000 |
| | Reliability-in-the-small | | | 0.0015 | 0.0018 | NA | 0.8573 |
| Intermediate-term forecasts: Brier score | | | | | | | |
| t=1 year | | 0.0048 | 0.25 | **0.0387** | **0.0652** | **0.0389** | NA |
| t=1.5 years | | 0.0048 | 0.25 | **0.3274** | **0.1258** | **0.0606** | NA |
| t=2 years | | 0.0048 | 0.25 | **0.0012** | **0.0012** | **0.0213** | NA |

Table 8 reports the Brier score (Brier, 1950) and its decomposition (Murphy, 1973) for the survival estimates generated by the models out of sample. A naïve model generating random forecasts of 0.5 has a Brier score of 0.25. The benchmark Brier score of each model was derived by setting the probability survival estimate for each holdout observation equal to the survival proportion observed in the estimation period. The model without SR (model 1) and the model with SR (model 2) were employed to form short- and intermediate-term forecasts for all holdout observations. The model with SR and CW (model 4) was used to generate short-term estimates while the models with SR and OL (model 3) were employed to form intermediate-term forecasts at six month intervals.

## Appendix A

As indicated in section 5.1, the Cox's dynamic hazard model of migration outcome *s* for state *m* at time *t* can be expressed as:

$$h_{s,m}(t, Z, Z(t)) \;=\; h_{s(0)}(t)\;\exp[Z_{\mathrm{j}}^{\mathrm{m}}\beta_{\mathrm{j}} + Z_{p}^{m}(\mathrm{t})\beta_{\mathrm{p}}\;]$$

The likelihood $L_{t_m,s}^{m}$ that state *m* experiences an event outcome *s* at time $t_m$ is calculated as follow:

$$L_{t_m,s}^{m} = \frac{\exp\left(\beta_j Z_j^m + \beta_p Z_p^m(t_m)\right)}{\sum_{i \in R(t_m,s)} \exp\left(\beta_j Z_j^i + \beta_p Z_p^i(t_m)\right)} \qquad \text{(A.1)}$$

Where: *i* represents a rating observation in the risk set formed at event time $t_m$ for event outcome *s*, $R(t_{m,s})$.

The time-varying covariate value $Z_p^m(t_m)$ used in the estimation process was updated to the most recent value as an event of interest occurred. Rating observation *i* appearing in different risk sets *R(t,s)* will carry different values of the time-varying covariates $Z_p^i(t)$ updated at various event times *t* when those risk sets were formed.

Taking the product of the likelihoods, for all states *m* that experienced event outcome *s*, across all event times $t_m$ observed in the estimation sample gives the partial likelihood, *PL*, as follow:

$$PL = \prod_{m=1}^{n_S} L_{t_m,s}^{m} = \prod_{m=1}^{n_S}\left[\frac{\exp\left(\beta_j Z_j^m + \beta_p Z_p^m(t_m)\right)}{\sum_{i \in R(t_m,s)} \exp\left(\beta_j Z_j^i + \beta_p Z_p^i(t_m)\right)}\right] \qquad \text{(A.2)}$$

Where: $n_s$ is the number of events of outcome *s* observed in the estimation sample.

The vectors of the estimated coefficients $\hat{\beta}_p$ and $\hat{\beta}_j$ can be obtained in the absence of knowledge of the baseline hazard $hs_{(0)}(t)$ by maximizing the full partial likelihood in Equation (A.2) (Kalbfleisch and Prentice, 1980).

The baseline hazard $hs_{(0)}(t)$ is not needed in the estimation process but is required to estimate the hazard of a future event. In the presence of the time-varying covariates $Z_p(t)$ the proportionality assumption of the conventional Cox's hazard model (Cox, 1972) does not hold and the baseline hazard $hs_{(0)}(t)$ cannot be extracted from the Cox's regression results. Estimating the baseline hazard function $hs_{(0)}(t)$ and forming the hazard of a future event from the dynamic Cox's hazard model with time-varying covariates $Z_p(t)$ is a challenging task.

This study uses the method proposed by Andersen (1992) and adopts the SAS codes published in a medical study by Chen, Yen, Wu, Liao, Liou, Kuo, and Chen (2005) to estimate the integrated base line hazard. Given the vectors of the coefficients $\hat{\beta}_p$ and $\hat{\beta}_j$ estimated in equation (A.2), the integrated baseline hazard $H_{s,(0)}(t)$ can be estimated as follow.

$$\hat{H}_{s,(0)}(t) = \sum_{t_m \le t} \frac{D_{m,s}}{\sum_{i \in R(t_m,s)} \exp\left(\hat{\beta}_j Z_j^i + \hat{\beta}_p Z_p^i(t_m)\right)} \tag{A.3}$$

Where: $D_{m,s}$ is an indicator for whether an event type $s$ occurred to state $m$ at time $t_m$ within the interval [0, t].

The integrated baseline hazard function $H_{s,(0)}(t)$ can also be estimated as a step function discontinued at event time $t_m$ (Chen et al., 2005).

$$H_{s(0)}(t) = \sum_{t_m \in t} [h_{s(0)}(t_{m-1})(t_m - t_{m-1})] \tag{A.4}$$

The estimated baseline hazard function at time $t$, $\hat{h}_{s(0)}(t)$, can then be derived from equations (A.3) and (A.4).

The estimated hazard of an event type $s$ for holdout state $q$ at time $t$ can be estimated using state $q$'s actual covariate vector $Z_j^q$ and $Z_p^q(t)$[29], the estimated baseline hazard function $\hat{h}_{s(0)}(t)$, and the estimated coefficient vector $\hat{\beta}_p$ and $\hat{\beta}_j$:

$$\hat{h}_{s,q}(t, Z, Z(t)) \ = \ \hat{h}_{s(0)}(t) \ \exp[Z_{\mathrm{j}}^{\mathrm{q}}\hat{\beta}_{\mathrm{j}} + Z_p^{\mathrm{q}}(\mathrm{t})\hat{\beta}_{\mathrm{p}} \ ] \tag{A.5}$$

The predicted survival function of holdout state $q$ at time $t$ can be estimated as:

$$\hat{S}_{s,q}(t, Z, Z(t)) = \exp - [\sum \hat{h}_{s,q}(t, Z, Z(t))] \tag{A.6}$$

[29] At the start of holdout rating state $q$, the subsequent migration time and the changes in macro-economic and political conditions over its survival duration are unknown. It is not possible to frequently update the values of the time-varying covariates $Z_p^q(t)$ over raring *q's* survival duration as only information up to the commencement of state $q$ is available. The values of the time-varying covariates $Z_p^q(t)$ used to form the predicted hazard for holdout observation $q$ were therefore measured at its beginning.

## Appendix B

The Brier score (Brier, 1950) was used to assess the predictive accuracy of probability survival estimates formed in Equation (A.6). The Brier score of survival estimates generated by the model for the migration outcome *s* at time *t*, $B_{s,t}$, is defined as:

$$B_{s,t} = \frac{\sum_{q=1}^{N_t} [f_{s,q,t} - d_{s,q})^2]}{N_t} \tag{A.7}$$

Where:

$f_{s,q,t} = \hat{S}_{s,q}(t, Z, Z(t))$, which is obtained from equation (A.6), is the probability forecast *f* that holdout state *q* will survive from the migration outcome *s* at forecast time *t* [30].

$d_{s,q}$ is the known outcome survival state *d* of holdout state *q*. If holdout state *q* survives from the event type *s*, $d_{s,q}$ = 1, and if holdout state *q* experienced the event of type *s*, $d_{s,q}$ = 0.

$N_t$ , or *N* for short, is the number of observations in the holdout sample at forecast time *t* , which is also the number of estimates at forecast time *t*.

The Murphy decomposition (Murphy, 1973) of the Brier score $B_{s,t}$ at forecast time *t* is given as:

$$B_{s,t} = \underbrace{\bar{d}_{s,t}(1-\bar{d}_{s,t})}_{\text{Outcome index variance}} + \underbrace{\frac{1}{N_t}\sum_{j=1}^{J} N_t^j (f_{s,t}^j - \overline{d_{s,t}^j})^2}_{\text{Reliability-in-the-small (calibration)}} - \underbrace{\frac{1}{N_t}\sum_{j=1}^{J} N_t^j (\overline{d_{s,t}^j} - \bar{d}_{s,t})^2}_{\text{Murphy resolution (discrimination)}} \tag{A.8}$$

Where:

$\bar{d}_{s,t}$ , or $\bar{d}$ for short, is the overall mean survival index, or the survival base rate, in the holdout sample for the migration outcome *s* at time *t*.

$\bar{d}_{s,t}(1-\bar{d}_{s,t})$, or $\bar{d}(1-\bar{d})$ for short, is the outcome index variance or the average of the squared distances between the sample relative survival frequencies $\bar{d}$ and the *N* outcome status $d_q$ of the *N* observations *q* in the holdout sample.

[30] The notation was changed to provide a compact presentation of the formula in a form consistent with the literature review on the Brier score

$f_{s,t}^{j}$, or $f^{j}$ for short, is the *j*th vector forecast. $f^{j} = (f_1^{j}, ..., f_{N_j}^{j})$ (*j=1, 2, , J;* $N = \sum_{j=1}^{J} N_j$ )

$N_t^{j}$ or, $N^{j}$, is the total number of observations on which the vector forecast is $f_{s,t}^{j}$.

$\overline{d_{s,t}^{j}}$, or $\overline{d^{j}}$ for short, is the sample relative frequency of the $N_t^{j}$ observations on which the sub-collections of forecast $f_{s,t}^{j}$ is offered.

$\frac{1}{N_t}\sum_{j=1}^{J} N_t^{j}(f_{s,t}^{j} - \overline{d_{s,t}^{j}})^2$, or $\frac{1}{N}\sum_{j=1}^{J} N^{j}(f^{j} - \overline{d^{j}})^2$ for short, is the reliability-in-the-small or the weighted average of the squared distances between the *J* distinct vector forecasts $f^{j}$ and the relevant sample relative frequencies $\overline{d^{j}}(j = 1, ...., J)$.

$\frac{1}{N_t}\sum_{j=1}^{J} N_t^{j}(\overline{d_{s,t}^{j}} - \overline{d}_{s,t})^2$, or $\frac{1}{N}\sum_{j=1}^{J} N^{j}(\overline{d^{j}} - \overline{d})^2$ for short, is the Murphy resolution or the weighted average of the square distances between the sample relative frequencies for the *J* sub-collections of forecasts $\overline{d^{j}}$ and the sample relative frequencies for the entire collection of forecasts $\overline{d}$ .