CHAPTER III

Data and Methodology

3.1 Data

In this section, it describes the data collecting process of the dependent and independent variables used in this study. The sample contains issuer rating data, accounting data, and stock price data in G7 except US. This study has divided the data into two periods: the first is the period from 1997 – 2004 to be the in the sample for estimating period and the second period is from 2005 – 2007 to be out-of the sample for forecasting period.

3.3.1 Issuer Rating Data

Issuer ratings of each country are obtained form Reuters, which contains monthly corporate rating data used in this study from 1997 – 2007. Rather than using bond ratings, this study uses the issuer ratings instead. One of the advantages of using issuer ratings over individual issue ratings is that any issuer rating change will reflect change in the financial and business situations of firms, rather than changes in indenture provision. All issuer ratings used are classified into 17 categories: AAA, AA+, AA, AA-, A+, A, A-, and so on[1]. Table 1 presents the total number of ratings of entire data set.

3.3.2 Independent Variables

In order to construct explanatory variables at the individual firm level, the DataStream is the source of data. Table 2 summarizes the descriptive static of explanatory variabls. The specific variables employed from Altman (2004) are WC/TA (net working capital to total assets), EBIT/TA (earing before interests and

---

[1] In order to have a reasonable number of observations in each rating class, the issuer rating classes CCC, CC, and C are combined to a single rating class CCC/CC.

taxes to total assets), ME/TL (market value of equities to total liabilities), and Age. In order to increase the effectiveness of the EBIT/TA and ME/TL variables in the logit model estimation, these variables are log-transformed as follow: -ln(1 -EBIT/TA)amd 1 + ln(ME/TL). The WC/TA variable is a proxy for the short-term liquidity of a firm. The EBIT/TA, and ME/TL variables proxy for current and future profitabilities, respectively. All of the Altman's variables should have positive relationships with upgrade, and negative relationships with downgrade.

The specific variables employed from Blume (1998) are OPTMAR, LTD/TA, TD/TA, Beta, and SE. Operating profit margin, OPTMAR, is a measurement of profitability. It should be positively related to improvement in credit ratings. The LTD/TA, long-term debt to total assets, and TD/TA, total debt to total assets, are proxy for the leverage. These two accounting ratios should be positively related to deteriorate in credit ratings. The next to market variables employed are Beta and SE are estimated by using at least 200 daily returns The expected signs on these two measures of risk are negative.

The variables employed from Shumway (2001) are NI/TA (net income to total assets), TL/TA (total liabilities to total assets). Both ratios are constructed as measures of profitability and leverage. Campbell et al. also found better explanatory power when they measured the equity component of total assets at market value by adding the book value of liabilities to the market value of equities. They are called NI/MTA (net income to market valued total assets) and TL/MTA (total liabilities to market valued total assets). A measure of liquidity, CA/MTA (cash and short-term assets to market value of its assets), is constructed. Market-based variables are added into this study. Diff-Ret is measured as monthly log excess return on each firm's equity relative to the index, Rel-Size is constructed as the logarithm of the ratio of the market

capitalization of a firm at the end of the year to the total market capitalization, and Sigma is the standard deviation of each firm's daily stock return over the past three months. The variable CA/CL (current assets to current liabilities) from Zmijewski (1984) is also added to measure liquidity.

In addition, this study examines the standard of rating agency whether it has lenient or stringent standard when assigning ratings. Therefore, time dummies are used to be measurements, but they do not distinguish the firm quality itself and effects of business cycle. Hence, I utilize four measures of business cycle indicators. The first two are indicator of the state of the economy that reflects by movements in level of employment, EMPLOY, and industrial production, PROD. The hypothesis is consistent with the perception that agency is too aggressive in downgrading ratings during bad economic times. The third and forth business cycle indicators do seek to capture both ups and downs in economic activity. In particular, I use the output ratio, OUTPUT, defined as the ratio of actual real GDP and potential GDP. The output ratio is a measure of how economics conditions relatively perform compared to the sustainable of economic activity. The other business cycle measure I consider is, unlike the output ratio, a discrete-valued indicator of the relative rate of current real GDP growth. It is defined as follows. The histogram of annual real GDP growth rates for entire sample period is constructed. If the current observation of annual growth falls into the lower third of this distribution, the indicator is assigned a value of -1, a 0 if it falls in the middle third and a 1 if it falls in the upper third. This indicator is employed from Nickell et al. (2000).

3.2 Research Hypotheses

To investigate which models is the best in predicting rating change and whether the rating agency has weaker standard in assigning rating, the following hypotheses will be empirically examined.

Hazard rate model incorporates time-varying covariates, or explanatory variables that change with time, unlike static model that does not concern time-varying. In the fact that firms change through time, therefore, hazard rate model should predict the rating change more precise than the static model.

Hypothesis 1: The hazard rate model can predict a rating change precisely than the static model.

According to the distress failure of giant companies such as Enron, WorldCom, and Parmalat even though they had investment rates they went bankrupt within a few months. Therefore the agency might have weaker standard in assigning ratings.

Hypothesis 2: The rating agency has weaker standard in assigning issuer ratings.

3.3 Methodology

This study investigates the precise performances on predicting issuer ratings change of static and hazard rate models and standard of rating agency S&P when it assigned ratings.

3.3.1 Predicting Issuer Ratings by Static and Hazard Rate Models

In this study, I examine the correct predicting performances of S&P issuer ratings between static and hazard rate models. In these estimations, the data from period 1997 through 2004 are used. Both models apply binary logit technique to

estimate parameters, which are based on the cumulative distribution function for the logistic distribution that is given by:

$$P_{t-1}(y_{i,t} = 1 | x_{i,t-1},\beta) = e^{x'\beta} / (1 + e^{x'\beta}) \quad (1)$$

where $y_{i,t}$ is an indicator that equals one if the firm's rating is changed in year t, and $x_{i,t-1}$ is a vector of explanatory variables known at the end of the previous year. The equation 1 can be written as follow:

$$Z_{it} = \beta X_{it-1} + \varepsilon_{t-1} \quad (2)$$

where Z is an latent variable which is one if rating is changes and zero if rating does not change, β is an estimated covariate, and X is a set of explanatory variable. This study uses five sets of explanatory variables. The first set is employed from the Altman's model which contains WC/TA, EBIT/TA, ME/TL, and Age. The second is employed from the Blume's model which contains OPTMAR, LTD/TA, TD/TA, Beta, and SE. The third set is from the Shumway's model is NI/TA, TL/TA, Dif-Ret, Rel-Size, and Sigma. The forth is from Campbell (2005) by adding MB and Price from Shumway, and the last set is added CA/CL is from the Zmijewski's model.

Even though both two static and hazard rate models apply logit technique they are different. Since static model uses the number of firm years in the data to estimate coefficients while hazard rate model uses each firm's entire event span is one observation to estimate coefficients. Static model, for instance, considers firm one which has a rating A in year 1998 and it is changed to A- in the interim of year 2000 and firm two which has a rating A in year 1998 and its rating does not change. Therefore, Z of year 2000 of firm one is set to one since its rating is changed in the period of year 2000, and its explanatory variables are from the data at the end of year 1999 while Z of year 2000 of firm two is set to zero because its rating does not change and its explanatory variables are the same as firm one by using the data at the end of

year 1999 in estimating. Hazard rate model uses all available data in estimating so its data used include data in year 1998 and 1999 too. They are used by setting Z of year 1998 and 1999 of both firms as zero because their ratings are A. Hence, the explanatory variables used in hazard rate model are not only lagged one year data but also all available data that those firms have.

3.3.2 Testing Rating Accuracy

Various rating methodologies and credit risk modeling approaches have been developed to test the accuracy for finding out the best model in forecasting ratings. The important reason of this testing stems from the fact that rating models of poor quality could lead investors to making wrong decisions. In this section, I use the estimated covariates received from estimating periods to test the power of predictions by using the data from 2005 through 2007. The receiver operating characteristic, ROC, is used to test the accuracy. I briefly explain the concept of an ROC curve.

Generally, the predicting performance of any models is based on the cutoff value C because if the cutoff value is changed the correct percentages on predicting is also changed. That means the change in cutoff value provides four possible results in predicting. If the rating score is below the cutoff value and the rating changes subsequently, the predicting is correct. Otherwise the predicting wrongly classified a non-change as a change. If the rating score is above the cutoff value and the rating does not change, the classification is correct. Otherwise, a change is incorrectly assigned to the non-change group. The ROC curve is a plot of hit rate, HR(C), versus false alarm rate, FAR(C). HR(C) is the number of ratings change predicted correctly with the cutoff value C to the total number of ratings change in the sample. FAR(C) is the number of non-change ratings that are classified incorrectly as changed ratings by using the cutoff value C to the number of non-change ratings in the sample. The rating

model's performance is better the steeper the ROC curve is at the left end and the closer the ROC curve's position is to the point (0,1). Besides, I calculate the areas under the ROC curves. The more areas under the curves mean the more accuracy of the model.

3.3.3 Measuring Standard of Rating Agency by Ordered Logit Model

Ratings are by their nature qualitative, discrete-valued indicators of creditworthiness. They also have a nature ordering, with AAA best, AA+ next best and so on. I therefore use of the ordered logit model in the empirical analysis.

The ordered logit model can be described as follow. Let Z_{it} be the rating of firm i at time t and X_{it} a vector of explanatory variable at time t that influence the determination of firm i's rating. Z_{it} is an integer-valued variable, which equals to 17 if AAA, 16 if AA+, and so on. The formula is as follows:

$$Z_{it} = \beta X_{ti} + \varepsilon_{it} \quad (3)$$

This equation looks like the equation 2 but they are different in the term of explanatory variables and the value of Z. The five sets of explanatory variables are used. The first set is employed from the Altman's model, the second is employed from the Blume's model, the third set is from the Shumway's model, the forth is from the Campbell's model, and the last set is from the Zmijewski's model, like the predicting issuer ratings change approach. Moreover, this part uses not only these accounting variables but also uses time dummies which measures the standard of rating agency and macro-variables, EMPLOY, PROD, OUTPUT, and GDP which measures the effect of business cycles. The time dummy of year 2000 is not included to the test because it is treated as a base-year. These time dummies will indicate the standard of agency on each year, the values of them will show the trend of the standard whether it is more stringent or more lenient. The macro variables are used

because issuer ratings of firms might be depended on the business cycle. If that time the economy is in an expansion period, firms will tend to receive high ratings, but if the economy is in a recession period, firms will tend to receive low ratings. Therefore, these factors have to be controlled. The sample of this part is listed and rated firms which existed during 2000 through 2007 since the size of sample have to be controlled to measure agency's standard in each year. To measure the standard of agency rating is to look at the coefficient values of each time variables, lower values on the coefficient implies more stringent grading standard but higher values on the coefficient implies more lenient standard of agency.