An Analysis of Peer Effects in the Credit Rating Market

A Dissertation presented

by

Bo He

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Economics

(Industry Organization)

Stony Brook University

May 2015

Stony Brook University

The Graduate School

Bo He

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Yiyi Zhou, Ph.D - Dissertation Advisor
Assistant Professor, Department of Economics
Stony Brook University

Sandro Brusco, Ph.D - Chairperson of Defense
Professor and Chair, Department of Economics
Stony Brook University

Mark R. Montgomery, Ph.D - Committee Member
Professor, Department of Economics
Stony Brook University

Xinyun Chen, Ph.D - Outside Committee Member
Assistant Professor, Department of Applied Math and Statistics
Stony Brook University

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

An Analysis of Peer Effects in the Credit Rating Market

by

Bo He

Doctor of Philosophy

in

Economics

(Industrial Organization)

Stony Brook University

2015

This paper empirically examines the peer effects among credit rating agencies (CRA) and the credit rating shopping behavior by bond issuers in the sub-prime mortgage market. The recent sub-prime mortgage crisis has raised the controversy that CRA's rating decisions might not be independent. For example, Moody's rating decisions might be affected by S&P, and vice versa. My studies analyzes the peer effects on Moody's and S&P's rating decisions, while taking into account the selection of CRAs by bond issuers. This selection process is also called credit rating shopping. At the issuance of mortgage-backed securities, bond issuers solicit ratings from CRAs. Only the selected CRAs will make their ratings public and get paid by bond issuers. Since both the credit rating shopping and the peer effects are likely to inflate ratings, ignoring the selection stage in the estimation will lead to upward bias of peer effects. The model is estimated using the data of sub-prime mortgage-backed securities from Jan. 2004 to Oct. 2008.

My studies shows the following findings: First, there is robust evidence of the peer effects on Moody's and S&P's rating decisions. The peer effects

on S&P are stronger than on Moody's. Second, the selection process does not significantly affect CRAs' rating decisions. Third, the peer effects are small for AAA bonds, then increase significantly for medium-rating bonds, and then decline again for lower-rating bonds. Fourth, Moody's and S&P's downgrade probabilities given its peer's rating one notch lower is much higher than their upgrade probabilities given its peer's rating one notch higher at all rating levels. Fifth, choosing two agencies has complementary effect as compared to choosing only one agency. Last, the downgrading probability at investment grade is not lower than the ones around investment grade. In other words, whether a bond is at the investment grade does not affect an agency's downgrading decision.

Keywords: Peer effects, Selection Bias, Mortgage-backed Securities

Table of Contents

Contents

List of Tables

List of Tables

List of Figures

List of Figures

Acknowledgements

Seven years have passed away since my very first step in US till this moment of writing the acknowledgement of my dissertation. It is a long journey for me. Dr. Hugo Benitez-Silva said in a graduation ceremony that "You will never know where is the finishing line till you have crossed it." The seven year's PhD life taught me lots of things. I found so many things in common between PhD research and real life: accept the failures humbly, analyze the problems wisely, start all over again boldly, and fight the battles confidently in all circumstances.

First of all, I would like to express my special thanks to my advisor, Dr. Sandro Brusco, who is always there to support, inspire, and encourage. His wisdom and sincere advices for me have lightened up my dark days and guided me to become more confidence and independent through out my PhD researches. I also want to thank Dr. Yiyi Zhou, who has provided tremendous helps and valuable insights for my thesis. She introduced me into this amazing empirical IO field, and taught me all those modeling and estimation skills in various IO topics. I enjoyed her IO course so much and took it twice. Sandro and Yiyi are not only role models who guided me in research, but also wonderful mentors that I truly appreciate. I would like to thank Dr. Wei Tan, who supervised me when I first started my PhD thesis. He has trained me to be a critical thinker and a better data analyst. I would also like to thank my committee member, Dr. Mark Montgomery, and my department director, Dr. Hugo Benitez-Silva, for their valuable suggestion whenever I need it.

Along with these great professors, I am also grateful for all my classmates and senior students. I present my thanks to Yajun Wang, E Yang, Siyu Qi, Yi Zhang, Jieruo Liu, Yoobim Kim, Xingzhi Yu, Qi Sun, Maryam Mahtab. We shared our happiness and frustrations, and we inspired and encouraged each other always. I appreciate the friendship with all of them for the great moments during my PhD, and they all amazing economist and my life-long friends.

I also want to thank my friends outside the department, especially Yiman Zhang, Luo Min, Liu Wei, Dongdong Zhou, Meng Yuan. Your love, humor, and comforting words have supported me through all the hard times.

In the end, my greatest thanks go to my parents for their unconditionally love and sacrifice. They might never understand my work, but they have always been my strongest backbones. I am forever grateful for the freedom they give me to explore the world, and to do what I want.

Lastly, I want to thank the seven years of PhD life. I will always cherish this experience, and keep faith, hope and love in my heart to start my next journey.

Chapter 1

1 Introduction of Credit Ratings

The quality of credit ratings is crucial to the stability and efficiency of the global financial system, since credit rating is widely used by market participants and regulators. Issuers seek for ratings not only because it is a requirement for asset securitization, but also because ratings can improve the marketability and pricing of securities. Investors rely on ratings to learn about professional view on the credit worthiness of securities and make investment decisions accordingly. Some regulatory constraints are also based on ratings. For example, certain institutional investors, such as pension funds and insurance companies, are forced to hold investment grade bonds by regulations; the Basel II Accord recommends a bank's required amount of capital to be calculated based on the ratings of the securities owned by that bank. Therefore, if credit ratings are not accurate and misrepresent the credit risk, regulations and investment practice based on ratings will be harmed.

The sub-prime mortgage crisis in 2008 brought the credit rating agencies (CRAs) into public's and regulators' attention. Credit rating agencies have been criticized for causing this crisis as their inflated ratings resulted in the rapid growth of the sub-prime mortgage lending. According to Fitch Rating (2007), around 60% of structured products[1] were AAA-rated (Nelson et al. (2012)). Meanwhile, Moody's assigned AAA-ratings to more than half of its structured products. Such high ratings swelled the market's confidence, and led to a spiral in structured products. One prominent sector in structured products which gained phenomenal increase is the sub-prime mortgage-backed securities. As the trigger of the financial turmoil, the origination of sub-prime mortgages rose from $65 billion in 1995 [2] to approximately $600 billion in 2006[3]. The soaring rating demand yielded great profit for CRAs. It tripled Moody's profits between 2002 and 2006. However, this booming period ended around the second half of 2006 and early 2007. Followed

[1] The main type of structured products include asset-backed securities (ABS), mortgage-backed securities (MBS), home equity loan securities (HELS),collateralized debt obligations (CDO), collateralized bond obligations (CBO), collateralized loan obligations (CLO), and collateralized mortgage obligation (CMO).

[2] Ashcraft and Schuermann (2008)

[3] Chomsisengphet and Pennington-Cross (2006)

by a subsequent jump in the sub-prime mortgage delinquencies, credit rating agencies substantially downgraded the ratings of a large amount of sub-prime mortgage-backed securities. Such downgrading wave concentrated among AAA-rated tranches. Among the 36346 tranches downgraded by Moody's in 2007 and 2008, nearly one third of them carried AAA ratings (Benmelech and Dlugosz (2010)). The large magnitudes of downgrades indicates that the ratings were likely inaccurate when they were first issued, and were not updated in a timely manner.

The Financial Crisis Inquiry Commission reported in Jan. 2011 that:" The three credit rating agencies were key enablers of the financial meltdown. The mortgage-related securities at the heart of the crisis could not have been marketed and sold without their seal of approval. Investors relied on them, often blindly. In some cases, they were obligated to use them, or regulatory capital standards were hinged on them. This crisis could not have happened without the rating agencies. Their ratings helped the market soar and their downgrades through 2007 and 2008 wreaked havoc across markets and firms."[4]

There are several explanations for CRAs' failure of issuing objective ratings to structured products. First, CRAs face conflicts of interests. They are paid by bond issuers whose products they are rating. Issuers can shop for ratings. A CRA is paid only when the credit rating is issued. If a issuer is unsatisfied about the rating, he may solicit another CRA. Therefore, CRAs have the incentive to lower rating standards and favor issuers in their competition for business. Cantor and Packer (1997) researched on corporate bond ratings and uncovered evidence that the credit rating agencies with lower market share rate corporate bonds more generously. Tan and Wang (2008) found in the sub-prime mortgage market that the more rating agencies rate a bond the less accurate those ratings are. Despite CRAs' financial motive to satisfy issuers with higher ratings, assigning accurate ratings is important for them to maintain good reputation for future business. A report by S&P to SEC in 2002 states that " the ongoing value of S&P's credit ratings business is wholly dependent on continued market confidence in the credibility and reliability of its credit ratings ".[5] Thus, we have the reasons to believe that CRAs' behavior is disciplined by their reputation concerns, although

[4]FCIC Financial Report Conclusions, Jan. 2011

[5]Standard & Poor's Rating Service, US Securities and Exchange Commission Public Hearing - November 15, 2002 Role and Function of Credit Rating Agencies in the US Securities Markets. http://www.sec.gov/news/extra/crdrate/standardpoorts.htm

the recent rating collapse in the financial crisis demonstrates that reputation is not a strong enough mechanism.

Second, CRAs face a trade-off between rating accuracy and stability. Rating accuracy refers to the correlation between ratings and risk of default. Rating stability refers to the frequency and magnitude of rating changes. An early survey conducted by Association for Financial Professionals (2002) reveals that most respondents believe that CRAs react too slowly to changes in corporate credit quality. Investors prefer more timely rating updates, even at the cost of increased times of rating reversals (Edward and Herbert (2006)). There are two sources caused investor's perception of CRAs' rigidity on rating changes. One is the Through-the-Cycle rating methodology. It is well-believed and widely used by CRAs. This methodology focuses more on long term credit risk than on short term fluctuation. As a result, a rating might not represent the Point-in-Time risk of the security. The other source for rating stability is CRAs' conservative rating migration policy. When CRAs perceive credit quality changes, they normally do not adjust ratings immediately, but wait till such changes stabilize or exceed some threshold.

Other reasons for CRAs' inaccurate ratings include the precision of their models. Deven Sharma, president of S&P, admitted that the historical data they used and the assumption they made significantly underestimated the severity of what actually occurred.[6]

Among all the researches of CRAs inaccurate ratings, little is known about how rating agencies influence each other's rating decisions. For example, whether a bond rating update from one CRA will cause other CRAs to follow? Similarly, is the fact that a CRA still maintains the original rating after detecting the deterioration of bond credit quality because none of its competitors have taken action? Very few researches investigate such impact of one CRA's action on the others and the differential influence based on a CRA's market size. This limitation restricts insight for financial regulators and hinders effective oversight.

This paper aims to examine the peer effects and credit rating shopping in the credit rating market. The peer effect means that one rating agency's decision is affected by other agencies. The credit rating shopping occurs when bond issuers select which CRAs to work with at a bond's issuance time. Both of the peer effect and the credit rating shopping will lead to

[6] "Testimony before the Committee on Oversight and Government Reform, United States House of Representatives", Deven Sharma, October 22, 2008

rating inflations. Therefore, we need to take into account the selection of CRAs by bond issuers in the analysis of peer effect. Ignoring the selection process will make the estimates of peer effect upward biased. Our hypothesis is that CRAs tend to keep close rating with their peers. Especially, if one CRA's current rating on a bond is lower than its peer's, its probability of downgrading that bond will be lower, which will harm its rating accuracy.

A CRA's rating decision could be affected by other CRAs for at least two reasons. First, CRAs provide rating service to bond issuers who ultimately choose which CRA to hire. The competition for business makes CRAs cautious to issue a rating lower than its peer's. Second, when a CRA's rating turns out to be inaccurate, the reputation cost usually depends on the relative accuracy of its rating. Therefore, CRAs have incentives to conform to each other and give similar ratings to reduce reputation damage in case that their ratings turn out to be inaccurate (Bajari and Krainer (2004)). Based on a data set of 17899 sub-prime mortgage bonds and their credit ratings from Jan.2004 to Oct.2008, I established a two-stage model to include the selection of CRAs by bonds issuers in the first stage, and measured the peer effectd on their rating decisions in the second stage. The estimation result shows robust evidence of peer effects among CRAs.

The following sections of the paper are arranged as follows. I give an introduction of credit rating agencies in section 1 and the structure of credit rating market in section 2. I summarize the previous researches on credit rating agencies in section 3 literature review. Then I present the data and the data processing in section 4. Section 5 provides an overview of sample selection models and their available estimation methods. In section 6, I build up a two-stage model, analyze the identification of model parameters, and explain the estimation methods. Section 7 interprets the estimation results and examines the peer effect without considering the selection process as well as the marginal effect of peer's rating. The last section concludes the paper.

Chapter 2

2 Industry Background

2.1 Credit Rating Agencies and Mortgage-backed Securities

Standard and Poor's (S&P), Moody's, Fitch, and Dominion Tranche Rating Service (DBRS) are the only credit rating agencies that rated sub-prime ABS in the United States. The two largest CRAs S&P and Moody's control 80% of the global market share, and the "Big Three" rating agencies Moody's, S&P, and Fitch control 95% of the rating business (Alessi (2013)). The "Big Three" CRAs are US-based companies, while DBRS is the largest rating agency in Canada with other offices in New York, Chicago and London. CRAs specialize in evaluating the credit risk of securities, using proprietary models to assess the default risk of the collateral and evaluate the strength of the deal structure. Bond issuers may also reveal non-public information to credit rating agencies. The rating agencies assign a letter grade to a specific bond based on the perceived long-term credit risk at the time of assignment. They have similar letter grade systems. In general, credit rating grades go from the top credit quality of AAA to AA, A, BBB, BB, B, C, and down to D. Within each rating grade, rating agencies normally refine rating further into "+" and "-", or "1","2" and "3". To facilitate our analysis of rating changes, we define a uniform set of numeric grades in order to differentiate ratings. Our numeric grades start with 1, which corresponds to the AAA rating, so the higher the number, the higher the credit risk. Rating agencies also interpret rating grades differently. For example, Moodys assigns Aaa rating for bonds "of the highest quality, with minimal credit risk", while S&P presents AAA rating as an indication that the "obligors capacity to meet its financial commitment on the obligation is extremely strong." Despite slight differences in CRAs letter grades, it is a general practice that bonds with BBB- ratings [7] or above are "investment grade", and those below BBB- are "speculative grade". C or D ratings are normally for bonds already in default with, as Moody's phrases it, "little prospect for recovery of principal or interest." Table 1 shows the mapping from letter grades to numerical

[7] BBB- corresponds to 10 in our numeric grade system

grades.

The securitization process of mortgages goes in the following way. Bond issuers originate loans, or acquire a pool of loans and transfer the assets into a stand-alone entity, called Special Purpose Vehicle(SPV). Loan borrowers pay back their mortgage monthly, and the cash flows from the asset pool are re-distributed to create a number of mortgage-backed securities with different priority of receiving monthly payments and bearing losses, which is known as tranches. Tranches are categorized into senior, mezzanine and junior levels. They are paid sequentially according to their priorities. Senior tranches are paid before mezzanine tranches, and mezzanine tranches are paid before junior tranches. Therefore, senior tranches have higher credit quality than the overall underlying asset pool. If the underlying mortgage loans default, junior tranches will be the first one to bear the losses.

Before a bond can be issued, it need to acquire a rating. Unlike corporate bonds on which CRAs make unsolicited ratings, ABS only have solicited ratings. When a bond issuer solicits ratings from credit rating agencies, each of agencies provides a "shadow" rating. That is an indication of how it would rate the bond given various possible deal structures. The deal structure is ultimately determined in a process of negotiation between the issuer and the CRA. After obtaining the bids from CRAs, the issuer finalizes the deal structure and decides to make which CRA's rating public(Chu and Fackler (2013)). In some cases, an issuer only publishes one rating, whereas in other cases, an issuer publishes two or three ratings. Issuer pays CRAs a small amount of initial fee for the shadow rating. If a CRA is selected to publish its rating, it will charge extra fee which is larger than the initial amount. To save on cost, an issuer may approach only a subset of CRAs for shadow ratings, and thus pays fewer initial fees. This rating shopping process cannot be modeled based on our data set because only the published ratings are observed, and it is unknown that which CRAs have been solicited ratings from by issuers.

The peer effect among CRAs can possibly occur in the initial rating assignment and in the ratings adjustment during the lifetime of a bond. As mentioned above, the initial rating assignment is an interactive process between CRA and bond issuers. The issuers adjust the structure of the deal according to the feedbacks from CRAs in order to achieve desired ratings. Though the CRAs which have been solicited ratings from by the issuers do not directly communicate with each other, such interactive nature of initial rating assignment process allows CRAs to be aware of the decisions of other

CRAs which are rating the same bond. The process possibly conveys information among CRAs (Sun et al. (2013))in the initial rating assignment. After bonds are issued, a CRA can observe other agencies' ratings and then make its rating decision accordingly. Therefore, there is high chance that one CRA being affected by its peer's decision when monitoring the quality of the bond after the bond issuance.

2.2 Market Structure

The unique market structure of credit rating services differentiate it from other markets of regular goods and services. The key feature of the credit rating market is the lack of competition. Three big players, S&P, Moody's, and Fitch, dominate the market with more than 95% market share in total (Atkins (2008)). The estimated market shares[8] of the three firms are approximately 40%, 40%, and 15% respectively. DBRS, which is also included in our data set, has much smaller market share than the big three. From 2004 to 2006 in our sample period, S&P and Moody's each rated more than 94% of sub-prime MBS at origination, while Fitch and DBRS rated 56.5% and 10.9% of sub-prime MBS respectively. With these figures, the Herfindahl-Hirschman Index for this market is 3233, much larger than the threshold of 1800 for a highly concentrated market.

The highly concentrated credit rating market can be attributed to three barriers. The first significant barrier to entry to the market is Nationally Recognized Statistical Rating Organization (NRSRO) designation. NRSRO has been created by SEC in 1975. During the 25 years after its creation, SEC first designated the big three and then four additional CRAs as NRSRO. However, merger and acquisition of small CRAs caused the number of NRSROs to return to the original three by the end of 2000. In Sept. 2006, Congress passed the Credit Rating Agency Reform Act into law, which instructed the SEC to cease setting up barrier to entry. SEC responded by designating more CRAs as NRSRO. By the end of 2007, there are ten NRSROs in the credit rating industry. Though obtaining a NRSRO designation is not necessary for CRAs to operate, having a NRSRO designation put them at a significant advantages. Small CRAs without a NRSRO designation are easily ignored by most investors and issuers, and thus are likely to remain small-scale (White (2010)).

[8] Based on revenues or issues rated

Economy of scale is another barrier to entry. The big three CRAs have built up relationship with major issuers through previous business. Such relationship ensures steady business flows and revenue in the future. CRAs established the knowledge base of deal structure and underwriting practice through experience, which also facilitates their future business. Large business volume will enable a CRA to invest in building a stronger evaluation system and market surveillance infrastructure, and a better evaluation system generally means better products and more customers (Tan and Wang (2008)).

Reputation plays an important role in the credit rating market. First, the threat of loss of reputation motivates CRAs to be self-disciplined and due diligent. Second, it creates barrier to new entrants. Each CRA has its own rating standards and methodology, and it takes time for customers to understand and accept a new rating standards and methodology. A customer who is familiar with a CRA's rating standards can associate its bond rating to default probability approximately accurately. Such familiarity is established on years of observation of the CRA's ratings and corresponding bond performance. Therefore, it is difficult for a new CRA makes its rating standards accepted by investors.

Chapter 3

3 Literature Review

The previous studies on CRAs mainly focus on the following areas. The first area which has been widely analyzed is the impact of competition on the conflicts of interests of CRAs. On one hand, CRAs have the incentive to understate the credit risk to attract business for current profits. On the other hand, they try to protect their reputation for future business. In this area, Bolton et al. (2012) establish a theoretical model to examine the rating game in the presence of rating shopping behavior, and find that CRAs are more likely to inflate ratings when reputation costs are lower, and the ex-ante surplus and investor surplus are both higher with a monopoly CRA than in a duopoly competition. Nelson et al. (2012) model the trade-off between maintaining reputation and inflating ratings, and find similar result that competition inflates the ratings and reduces the expected welfare. Becker and Milbourn (2011) provide empirical support for this result, showing that competition will lower the rating quality. Based on the bond rating data from 1998 to 2006, they found that the growth of Fitch's market share increases the overall credit ratings issued by Moody's and S&P, and the correlation between bond yields and ratings falls. It implies that ratings become less informative with increased competition. Among the early papers which generally investigate the trade-off between reputation and short-run profit from lying, Klein and Leffler (1981) also argue for the adverse effect of competition on maintaining reputations.

Though quite a few papers agree on the inappropriateness to introduce further competition into credit rating market by regulators, Bar-Isaac (2003) points out that the competition effect on quality is ambiguous and may be non-monotonic. He suggests that firm produces high quality only when the short-term cost of producing high rather than low quality is less than the difference between discounted value of high and low reputation. While the increased competition reduces the discounted value of high reputation on one hand, and increases the punishment of low reputation on the other hand, the competition can either enhance or hinder the reputation incentives for quality. Horner (2002) shows that competition increases the quality since the loss of reputation imposes threat of exit on firms. Xia (2012) provides empirical evidence that competition improves the rating quality. He compares S&P's

rating before and after the entry of an investor-paid credit rating agency, and finds that S&P's rating quality has been significantly improved after the entry of the new rating agency. But his study is the competition influence from an investor-paid rating agency, which is not the same as our interests on the competition among rating agencies which are all issuer-paid.

In their competition for business, CRAs differentiate issuers when assigning ratings. Mahlmann (2011) discovers that CRAs assign better ratings to firms with which they have longer relationship, though those firms do not have lower default rates. Compared to the corporate bonds, the structured finance products, such as mortgage-backed securities, have more complicated structures and cash flows. Thus, it becomes more difficult to detect CRAs' deviation from truthful ratings if such misconduct occurs. He et al. (2003) examine whether CRAs favor large issuers on the MBS market. They show evidence that CRAs' ratings are biased towards large issuers, because they can bring more business and revenue in the future. Such incentive to favor large issuers is even stronger in market boom period.

In all the aforementioned papers which model the competition, none of them consider the dynamic property of the credit rating market, and only consider competition in duopoly. Stefan (2013) fills the gap by using evolutionary game theory to analyze the interaction of CRAs over an infinite horizon in a competitive market with an arbitrary number of agencies. He points out the reason that Becker and Milbourn (2011) do not find an increase rating quality followed by the entry of Fitch is because the competition is not enough. There are such a few number of CRAs in the market that investors and issuers do not have sufficient number of alternatives, which results in very low reputation cost. By modeling the interaction among investors, CRAs and issuers, he solves for the critical number of CRAs for all the possible equilibria.

The second research area on CRAs is the trade-off between rating accuracy and stability. Investors value rating stability because volatile ratings are costly. The widely cited example of actions based on ratings which is costly to reverse involves the rating-based bond portfolio composition guidelines (Cantor and Christopher (2006)). To maintain the specific composition of portfolio based on ratings, fund managers have to purchase, sell and then repurchase the bonds if bond ratings changes frequently. Those reversed actions will generate large transaction costs which could have been avoided. CRAs claim that they update ratings based on long-term perspective and suppress the rating sensitivity to short-term fluctuation.

In terms of the rating horizon, ratings can be produced either on a Point-in-Time(PIT) or a Through-the-Cycle(TTC) basis. PIT method is a Merton-type model (based on Merton(1974)) and measures the default probability from short term risk. TTC method only focuses on persistent changes of the bond quality, which is in a long-term view. Carey and Hrycay (2001) describe the TTC method as rating in the bottom of the credit quality cycle (i.e. the default probability on short term). TTC imposes a stress scenario on the bond at the first time rating the bond. Thus, the initial bond rating under TTC method is generally lower than under PIT method, and TTC rating is insensitive to the credit quality cycle later. Gunter (2004) implements the TTC method by separating the permanent and cyclical components of default risk in a time-series setting. His paper concludes that TTC method has relatively low capability to predict default risk despite of its stability. John et al. (2013) extend and modify Gunter (2004)'s model, and compare the impact of TTC and PIT methods. They find that TTC method suffers from rating cliff effect due to its delay of rating changes. It also has inferior performance in predicting default relative to PIT method. In addition to the widely employed TTC method, another important source of bond stability is CRAs' prudent migration policy (i.e. "wait and see" policy). Using a credit scoring model, Edward and Herbert (2006) finds that rating migration is triggered only when the credit quality exceeds a threshold level of 1.25 notch steps. If it is triggered, ratings are only partially adjusted by 75%. This "wait and see" policy, along with TTC method, explains the slow adjustment of ratings by CRAs.

Other empirical studies on credit ratings include the analysis of the correlation between ratings and corporate default (Zhou (2001) and Jorion and Zhang (2007)), the assessment of impact of ratings on capital market (Gonzales et al. (2004)[9] and Followill and Martell (1997)), and the determinants of ratings and rating changes (Altman (2001) and Kamstra et al. (2001)). The classical methods used to forecast credit ratings include OLS regression, ordered probit model, unordered logit model, and the multivariate discriminate analysis (MDA). Ederington (1985) finds that the unordered logit and ordered probit outperform the OLS and MDA methods. Moreover, unordered logit achieves the best fit for in-sample estimation and ordered logit works best for out-of-sample prediction.

[9]Gonzales et al. (2004) find that market react stronger to downgrades than to upgrades. Downgrades have significant impact on stock prices while upgrade have no effect.

Among various researches of credit rating market, my paper focuses on peer effects among credit rating agencies. The peer effects have been widely studied in education and labor economics. For example, Ammermueller and Pischke (2009) estimate the peer effects for fourth graders in six European countries. Li et al. (2013) quantify the peer effects in a group lending program in India. From empirical point of view, Manski (2000) and Brock and Durlauf (2007) categorize peer effects into endogenous peer effects and contextual peer effects based on the channel through which peer effects operate. Endogenous peer effects capture that one's behavior can be directly affected by the behavior of his/her peers. Contextual peer effects refer to how the characteristics of a group can affect the behaviors of its members. Endogenous peer effects distinguish from contextual peer effects since the former give rise to "multiplier effects" through the feedback in member behaviors whereas the latter do not. In my analysis, the peer effects among CRAs refer to endogenous peer effects. That is how one agency's decision could be directly affected by the ratings of other agencies. CRAs tend to conform to their peers' ratings for two reasons. First, CRAs are selected and paid by bond issuers. Bond issuers tend to select CRAs which give more generous ratings (Faltin-Traeger (2009)). It would be difficult for a CRA to compete for business when issuing a lower rating than its peers. Second, a CRA's reputation cost of assigning an inaccurate rating depends on the relative accuracy of its rating (Sun et al. (2013)). The reputation cost would be lower if the CRA is wrong together with its peers than being wrong alone. Therefore, CRAs have the incentive to give similar ratings in order to reduce reputation cost in case that their ratings turn out to be inaccurate (Bajari and Krainer (2004)).

From methodology standpoint, the peer effect analysis generally adopts a game framework. The empirical studies on peer effects focus on the strategic interactions among players. It usually specifies a static game with incomplete information. The estimation of the game is carried out by a two-step approach developed by Bajari et al. (2013). They find strong peer effects among equity's analysis for NASDAQ stocks. In their two-step approach, they first solve a fixed-point problem to find out the equilibrium of each player's decision given his/her expectation on peers' decisions. Then they apply the equilibrium decision from the first step to recover the parameters in the second step based on the observed data.

With the retrospective of existing literature of credit rating market, I find very few empirical papers study CRAs' peer effect with the consideration of the bond issuers' credit rating shopping behavior (i.e. the selection process of CRAs). However, both the credit rating shopping and peer effect leads to rating inflation. Without considering the selection of CRAs in the first stage will make the estimation of peer effects up biased. One paper adds to this field of literature is Sun et al. (2013)'s studies on the strategic interaction in the credit rating market while explicitly accounting for the selection of CRAs by bond issuers. Their paper examines the peer effects at the initial rating assignment when CRAs provide their ratings simultaneously. My paper analyzes the peer effects on CRAs' rating change decisions after bond issuance, and controls for the quality change of bond using delinquency rates. In addition, I estimate the complementary/subsitutionary effect of choosing multiple agencies. In my paper, the selection process and peer effects are modeled in two stages. In the first stage, bond issuers choose which CRAs to work with. In the second stage, given the selected CRAs, the model examines whether the one CRA's rating will be affected by its peers' rating decisions. My analysis shows the existence of peer effects on Moody's and S&P's rating decisions. Moreover, it suggests the heterogeneity of peer effects in three dimensions: First, the magnitude of peer effects on S&P is greater than on Moody's. Second, the peer effects vary across ratings. The peer effects are small for AAA bonds, then increase significantly for medium-rating bonds, and then decline again for lower-rating bonds. Third, peer effects on downgrade probabilities are stronger than on upgrade probabilities. Either Moody's or S&P's downgrade probabilities given its peer's rating one notch lower is much higher than their upgrade probabilities given its peer's rating one notch higher at all rating levels. Last, the downgrading probability at investment grade is not lower than the ones around investment grade. In other words, whether a bond is at the investment grade does not affect an agency's downgrading decision. In contrast to Sun et al. (2013)'s studies, my analysis does not show significant selection effect on the rating decisions.

Chapter 4

4 Data

4.1 Data Description

The data set comprises 17889 sub-prime mortgage-backed securities issued from Jan. 2004 to Dec. 2006. The ratings are observable on a monthly basis till Oct. 2008. Each bond is identified by a unique Committee on Uniform Security Identification Procedures(CUSIP) code.[10] Bond characteristics and ratings are acquired from Intex Solution Inc., a major structured securities data supplier and related analytical software vendor. The delinquency rates of bonds at the deal level are extracted from Bloomberg from bond issuance date to Nov. 2010. The delinquency rates include 30-day, 60-day, 90-day, and 90-day plus delinquency rates. A bond is considered as default by practitioners if it has not been paid after 90 days. Table 2 shows two examples of bond observation in the data set. Insured bonds are excluded from the analysis, since they are all AAA bonds and do not have rating changes.

Before analyzing the peer effects among CRAs, the sufficiency of data has been checked from two perspectives - how many bonds are rated by multiple agencies and how often their rating changes are. Table 3 shows the number of bonds rated by different groups of CRAs. In my data, 44.66% of the bonds are rated by two agencies, 43.06% of the bonds are rated by three agencies, and 7.91% are rated by four agencies. Moody's and S&P jointly rated 38% of the bonds; Moody's, S&P, and Fitch jointly rated 39.11% of the bonds; Moody's, S&P, Fitch, and DBRS jointly rated 7.91% of the bonds. Therefore, there are enough number of bonds rated by multiple agencies for the peer effects analysis among those agencies.

Among the bonds rated by multiple CRAs, Table 4, Table 5, and Table 6 list how many of them have their ratings changed by one, two, three, and four CRAs respectively, as well as how many bonds that each CRA being the first one to change its rating. For example, among the bonds rated by both Moody's and S&P, there are 876 bonds whose ratings have been only changed by Moody's, 199 bonds whose rating have been only changed by S&P, and 2740 bonds whose rating have been changed by both Moody's and S&P. Among the bonds which have their ratings changed by both Moody's

[10] A standard security ID

and S&P, there are 1622 bonds which have their first rating change made by Moody's, 1099 bonds which have their first rating change made by S&P, and 19 bonds which have their first rating change made by both Moody's and S&P simultaneously.

Table 7 shows the number of rating changes that the four CRAs made on bonds respectively. Take Moody's for example. During the sample period, Moody's made no rating change on 7402 bonds, one rating change on 3434 bonds, twice of rating changes on 3002 bonds, three times of rating changes on 1730 bonds, four times of rating changes on 59 bonds, and five times of rating changes on 7 bonds. The bonds which have rating changes count for 52.65% of bonds rated by Moody's, 44.76% of bonds rated by S&P, 52.94% of bonds rated by Fitch, and 57.52% of bonds rated by DBRS.

Table 8 shows the market coverage of the four CRAs by the original balance of bonds rated and by the number of bonds rated respectively. In order to have the sum of four CRAs' market coverage being 100%, an indicator variable whether a bond is rated by a CRA is divided by the number of ratings on the bond. The market coverage of a CRA by original balance is the ratio of the sum of original balance of each bond rated by the CRA multiplied by the indicator variable over the total original balance of bonds in the market. Similarly, the market coverage of a CRA by number of bonds is the ratio of the sum of the indicator variable of each bond rated by the CRA over the total number of bonds in the market.

It shows that the market coverage by original balance of bonds rated by Moody's, S&P, Fitch and DBRS are 38%, 40%, 19%, and 3% from 2004 to 2006. The market coverage by the number of bonds rated are 37%, 42%, 18%, and 3% during the same period of time. According to their market coverage, Moody's and S&P are the largest two players in the market. They each accounts for around 40% of the market. Fitch has much smaller market coverage, less than 20%. DBRS only takes 3% of the market share. Therefore, my analysis of peer effect focuses on the interdependence between Moody's and S&P. Only bonds rated by Moody's or S&P or both are included in the analysis. With this criteria, I end up with 5914 bonds in the selection stage. Since the selection occurs at bond issuance date, the data used for selection stage estimation is cross-sectional. In the second stage, the ratings is used from each bond's issuance date to Oct. 2008. There are 40998 observations in the rating stage.

4.2 Data Analysis and Processing

Since the data set only has bond characteristics and rating changes, CRA characteristics were constructed based on the data. The constructed variables are CRA's market share of each issuer and the expected rating from a CRA based on bond's original support. The market of an issuer is defined as the total number of bonds issued by that issuer in a given time period. A CRA's market share of an issuer is thus defined as the number of bonds the CRA rated proportional to the total number of bonds issued by the issuer in that time period. When an issuer selects CRAs, the issuer would have an expected rating from that CRA. The expected rating could be based on bond's characteristics that affects the rating decision. Here I chose the original support. CRA has different ratings for different levels of original support. Therefore, if we want to use original support to infer a CRA's rating on the bond, we can quartile the original support of all the bonds up to the previous quarter of the bond's issuance date, and then calculate the average rating in each tile. The expected rating from that CRA is the average rating of the tile which the bond belongs to.

Table 9 provides an overview of variables used in the study. There are 5338 bonds rated by Moody's, 5887 bonds rated by S&P. The average rating of Moody's in the data set is 4.73, which is approximately "A1" in Moody's system. The average rating of S&P in the data set is 4.79, which is approximately "A+" in S&P's system. The standard deviation of Moody's and S&P's initial ratings are around 3.5 notches. The average balance of bonds at the issuance date is 57.55 million USD, while the average original support is 11.6%. The average coupon rate is 6.19%. 11% of the bonds are structured as fixed-rate. On average, there are about 15 tranches structured in one deal. In terms of market share, on average for all the issuers' market, Moody's takes 91% of the market share in last quarter before current business deal, while S&P takes 98% of the market share in last quarter before current business deal. Based on the original support, the average expected rating by Moody's and S&P are 4.75 and 4.38 respectively, corresponding to approximately "A1" in Moody's system, and "AA-" in S&P's system.

In the original data set, rating is observed on a monthly basis. I transformed it into a quarterly basis to increase the frequency of rating changes. Since my interest is in how Moody's rating is affected by S&P and vice versa, only bonds rated by both Moody's and S&P's are included in the second stage. The remaining sample used for the second stage consists of 40998 observations. Each observation has two ratings on them, and thus there are 81996 bond-CRA pairs.

As a time-varying variable, the delinquency rate is included in the rating equation to control for the quality change of bonds. However, there are two things need to be considered in using this variable. The first one is which delinquency rate to use. To measure the credit quality of a bond, I use the 90-day plus delinquency rate of the underlying mortgages. The most often used delinquency rates are the 30-day, 60-day, 90-day, and 90-day plus delinquency rates. The 30-day delinquency rate is the percentage of outstanding balance of delinquent loans in the past 30 days over the outstanding balance of all loans in a deal Gan and Mayer (2007). If a loan borrower misses consecutive two payment dates, then he will be accounted into the 60-day delinquency rate. The 90-day plus delinquency rate is defined as the sum of the percentage of loans that are grouped within the 90-days, REO[11], bankruptcy, and foreclosure delinquency buckets.[12]

The 30-day, 60-day, and 90-day delinquency rates are not cumulative but for current month. During the period of time of one or two months, a loan could become delinquent, and then current and then delinquent again. Each of the rates fluctuates without a monotonic trend since the bond issuance. Their changes are volatile. However, if a loan is delinquent for over consecutive three months, it has very small chance to become current again. Once the loan defaults, it will be excluded from the deal. Thus, the 90-day plus delinquency rate keeps a positive monotonic trend, and it shows more convincing results of the loan performance than other delinquency rates. Figure 1 to Figure 4 plot the changes of delinquency rates from 2004 to 2009. The 90-day plus delinquency rate looks more like an cumulative rate than 30-day, 60-day and 90-day delinquency rates. In comparing the loan performance, a cumulative delinquency rate is needed since it could show the process of how much of the asset pool goes delinquent or default during its lifetime. It is difficult to compare the quality change between two asset pool using non-cumulative delinquency rates. For example, a decrease of the 30-day delinquency rate does not necessarily mean the improvement of the quality of underlying assets. It could be that the mortgages are unpaid for more than

[11]REO is Real Estate Owned by the servicer. It is the percentage of all bank-owned property, except that taken in consideration of a default loan. These values are directly derived from values from the loan tapes.

[12]The definition of 90-day plus delinquency rate is from the Bloomberg Data Department of Fixed Income.

30 days and then fall into 60-day delinquent category. As shown in Figure 5, during the sub-prime mortgage crisis from 2008 to 2009, the 30-day and 60-day delinquency rates decreased but the 90-day delinquency rate surged rapidly. It indicates that those 30-day or 60-day delinquent loans are mostly become 90-day delinquent loans. The other way around, an increase of the 30-day delinquency rate does not necessarily mean the deterioration of the underlying assets. It could be that the 60-day or 90-day delinquent mortgages are partially paid and become 30-day delinquent. Figure 6 compares the four delinquency rates. It shows that only the 90-day plus delinquency rate has a positive monotonic trend as 90-day plus delinquent mortgages are unlikely to be paid back. Therefore, the increase of 90-day plus delinquency rate is a preferable indicator of the worsening of the underlying assets over other delinquency rates. Only the 90-day plus delinquency rate is used in my analysis to ensure the accuracy of results.

Another thing need to be considered when using delinquency rate is that the delinquency rate is at deal level. It is inconsistent with ratings and other bond characteristics which are at bond level. A deal includes multiple bonds, which are all supported by the same underlying asset pool. Those bonds are in different tranches and have different levels of risk. Junior tranche suffers the loss before senior tranche when the underlying loans default. Therefore, junior tranche carries higher risk than senior tranche and also has higher returns. The delinquency rate at deal level cannot capture the risk of individual bonds. One appropriate way to use the delinquency rate as a control variable for bond quality is to use tranche dummy $\times$ 90-day plus delinquency rate. The tranche dummy denotes which tranche the bond belongs to.

The model is estimated with and without the interaction term of tranche type and 90-day plus delinquency rate respectively, and compares the peer effect under these two specifications. The result shows that peer effect estimated without the delinquency rate is higher than the one estimated with the delinquency rate. In other words, delinquency rate partially explains the rating changes made by CRAs additional to the peer effects. Excluding the delinquency rate from the rating equation will cause the estimation of peer effects upward biased. More details is explained in section Estimation Results.

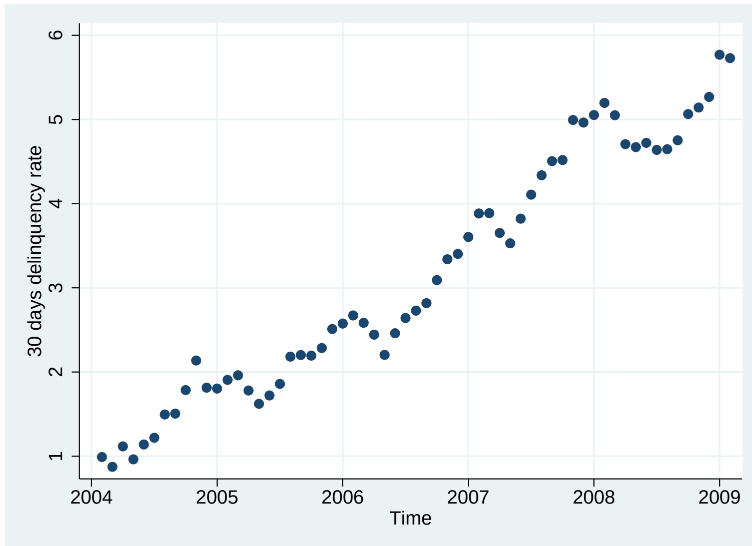

Figure 1: 30-day delinquency rate

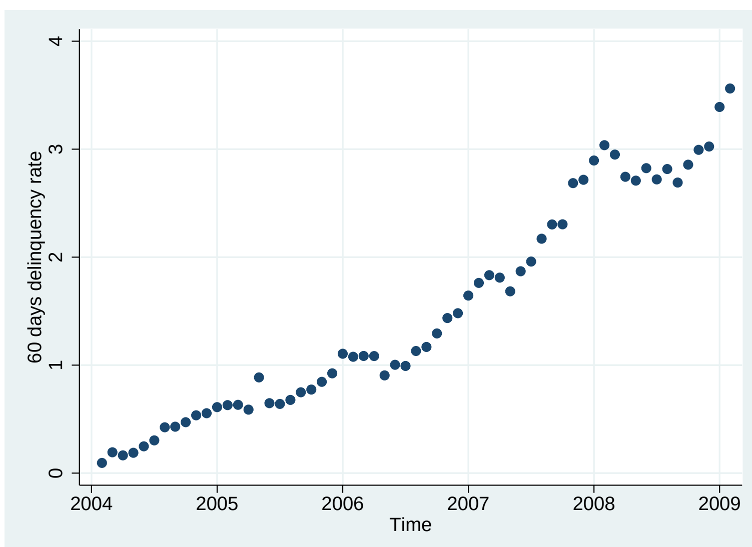

Figure 2: 60-day delinquency rate

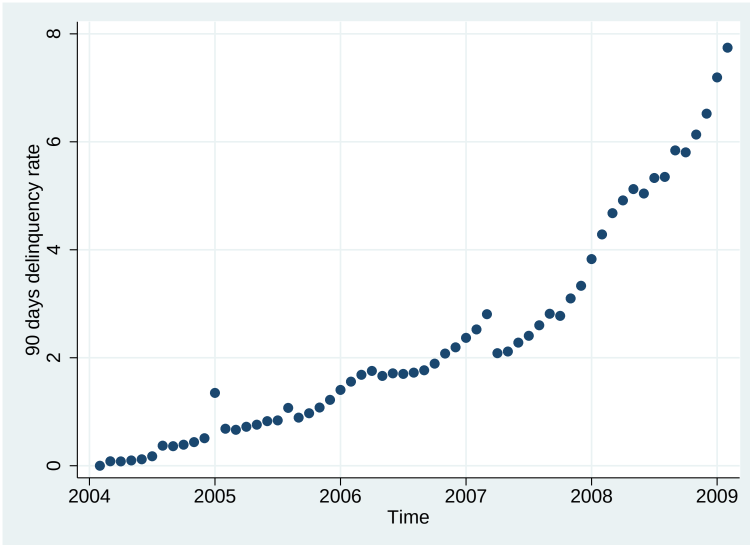

Figure 3: 90-day delinquency rate

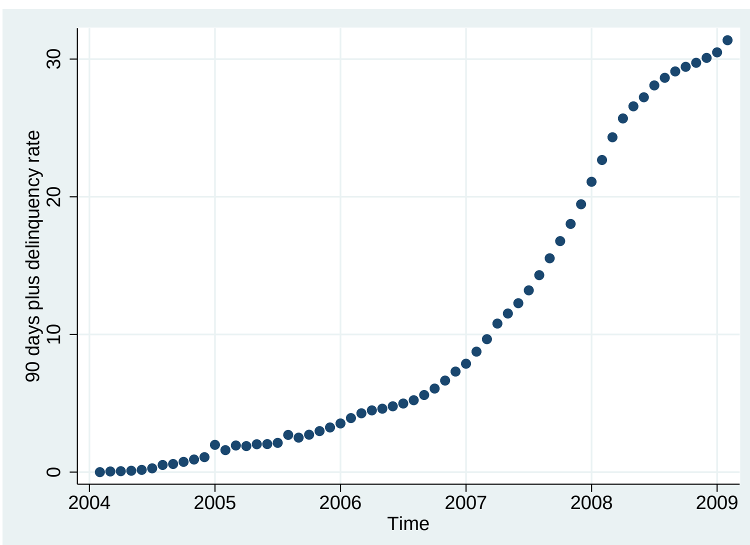

Figure 4: 90-day plus delinquency rate

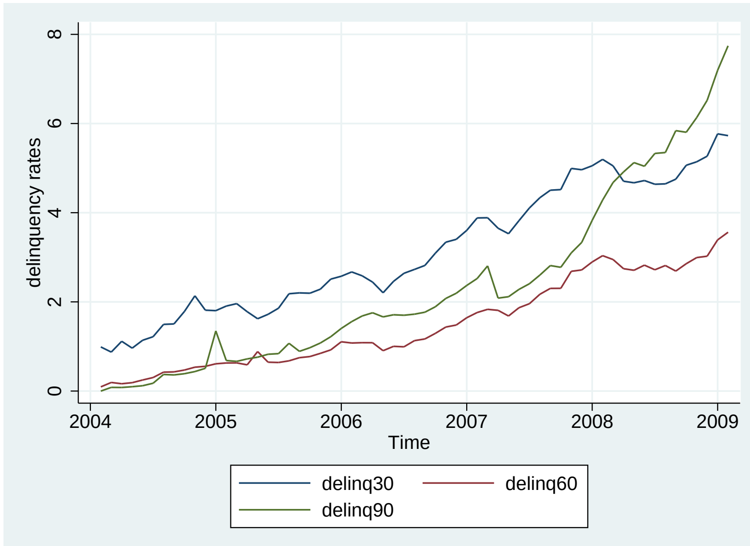

Figure 5: 30/60/90-day delinquency rates

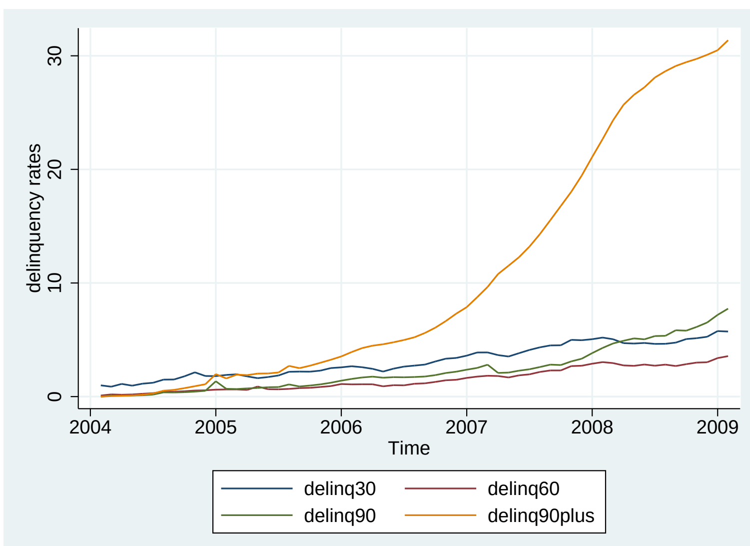

Figure 6: All delinquency rates

Chapter 5

5 An Overview of Sample Selection Models

This section provides an overview of sample selection models and their estimation methods that I studied for my model development. If a sample is not pure random and the usual estimators are applied, the estimates will be biased. A pure random sample is obtained by exogenous sampling [13], which is rare in reality. If instead a sample is based on values taken by a dependent variable, intentionally or unintentionally, parameter estimates may be inconsistent unless corrective measures are taken. Those samples are deemed as selected samples (Cameron and Trivedi (2005)). In this section, I start with introducing the bivariate sample selection model studied by Heckman (1979), and then extend it to the multinomial logit models with selection bias corrections studied by Lee (1983), Dubin and McFadden (1984), and Dahl (2002). Then I explain how I build up my selection model based on Gentzkow (2007) which allows consumers to choose multiple alternatives simultaneously. Lastly, I show how to apply the GHK method to estimate my model.

5.1 Sample Selection Model Estimated by Two-stage Regressions

We need to take into account the selection bias in estimating peer effects because we can only observe the ratings from agencies which are selected to rate the bonds. The selection of rating agencies by bond issuers depends in part on their relationships and issuers' expectation of agency's ratings. The complete model specification can be referred to the next section. A simplified sample selection model of credit rating agencies takes the following form.

Define the base utility of the issuer k for selecting CRA j to rate bond i to be

$$U_{i,j} = X_i\beta_1 + F_j\beta_2 + \xi_j + v_{ij}, \tag{1}$$

[13] Exogenous sampling occurs if the available sample is segmented into sub-samples based only on a set of exogenous variables, but not on the response variable (Cameron and Trivedi (2005)).

where X_i is a vector of observed characteristics of bond i; F_j is a vector of observed characteristics of CRA j; v_{ij} is the unobserved error. The second stage rating decision is modeled as

$$(R_{i,j,t}|S_{i,j} = 1, S_{i,j'} = 1) = \gamma R_{i,j',t-1} + \alpha X_i + \eta F_j + \epsilon_{i,j} \tag{2}$$

An OLS regression of $R_{i,j,t}$ on $R_{i,j',t}$, X_i and F_j alone using just the observed ratings from selected agencies leads to inconsistent estimation of coefficients unless the errors of the two stages are uncorrelated, that is $\sigma_{12} = 0$. Since errors of the two stages might be correlated, sample corrective measures need to be taken in the estimation. The rest of this section begins with the discussion of estimating bivariate sample selection model before focusing on the estimation of multinomial logit sample selection model.

A bivariate sample selection model comprises a decision equation that

$$y_1 = \begin{cases} 1 & \text{if } y_1^* > 0 \\ 0 & \text{if } y_1^* \leq 0 \end{cases} \tag{3}$$

and an outcome equation that

$$y_2 = \begin{cases} y_2^* & \text{if } y_1^* > 0 \\ - & \text{if } y_1^* \leq 0 \end{cases} \tag{4}$$

This model specifies that y_2 is observed only when $y_1^* > 0$. The latent variables y_1^* and y_2^* have the linear forms:

$$y_1^* = x_1^{'}\beta_1 + \varepsilon_1 \tag{5}$$

$$y_2^* = x_2^{'}\beta_2 + \varepsilon_2 \tag{6}$$

where ϵ_1 and ϵ_2 are correlated, with

$$(\varepsilon_1, \varepsilon_2) \sim N[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}] \tag{7}$$

The conditional truncated mean in the bivariate sample selection model is

$$\begin{aligned} E[y_2|x, y_1^* > 0] &= E[x_2^{'}\beta_2 + \varepsilon_2 | x_1^{'}\beta_1 + \varepsilon_1 > 0] \\ &= x_2^{'}\beta_2 + E[\varepsilon_2|\varepsilon_1 > -x_1^{'}\beta_1] \end{aligned} \tag{8}$$

where x denotes the union of x_1 and x_2. If the errors are independent then the second term becomes $E[\varepsilon_2] = 0$, and the OLS regression of y_2^* on x_2 will give a consistent estimate of β_2. However, any correlation between the two errors means that the mean is no longer $x_2^{'}\beta_2$ and we need to account for selection (Cameron and Trivedi (2005)). Heckman (1979) noted that if the errors $(\varepsilon_1, \varepsilon_2)$ are joint normal as in equation (7) then it implies that

$$\varepsilon_2 = \sigma_{12}\varepsilon_1 + \xi \tag{9}$$

where the random variable ξ is independent of ε_1. By using equation (9), equation (8) simplifies to

$$\begin{aligned} E[y_2|x, y_1^* > 0] &= x_2^{'}\beta_2 + E[(\sigma_{12}\varepsilon_1 + \xi)|\varepsilon_1 > -x_1^{'}\beta_1] \\ &= x_2^{'}\beta_2 + \sigma_{12}E[\varepsilon_1|\varepsilon_1 > -x_1^{'}\beta_1] \\ &= x_2^{'}\beta_2 + \sigma_{12}\lambda(x_1^{'}\beta_1) \end{aligned} \tag{10}$$

where $\lambda(z) = \phi(z)/\Phi(z)$ [14]. Estimation by maximum likelihood is straightforward given the above assumption on errors. One standard way to estimate a bivariate sample selection model is to use Heckman two-step estimator. Heckman's two-step procedure augments the OLS regression by an estimate of the omitted regressor $\lambda(x^{'}\beta)$. This term is called inverse Mills ratio. It is important to note that both the OLS standard errors and heteroskedasticity-robust standard errors reported from the second-stage regression are incorrect, the formulas for the correct standard errors are given in Heckman (1979), or to use the bootstrap.

Heckman's two-stage estimation method is only applicable to binary selection model for the selection bias correction. Lee (1983) and Dubin and McFadden (1984) extended Heckman's two-stage estimation to multinomial logit-based selection model. The difference between these two methods is their assumptions. Lee's method makes stronger assumption than Dubin and McFadden's, but avoids the risk of multi-collinearity present in the latter. Dahl (2002) provided a semi-parametric approach, whose variants also depend on the stronger precision arbitrage. Franois et al. (2004) used the Monte Carlo experiment to compare the advantages and shortcomings of these available methods in literature, and found that Dubin-McFadden and variants of Dahls models perform well, with relatively little efficiency losses

[14]See Page 566 of Cameron and Trivedi (2005) for the derivation of equation (10)

provided sample sizes are in line with micro-econometric contemporary practice. The Lee approach would only be warranted in very small samples. They also considered the case where the underlying selection process follows a polychotomous normal model, allowing correlations between alternatives. The multinomial selection bias correction methods perform well even in this case, with the flexible methods being again preferred.

I used my two-stage model in (1) and (2) as an example to illustrate how to use Lee and Dubin and McFadden's approaches and its variants to estimate a multinomial logit-based selection model. When the selectivity is modeled as a multinomial logit, the model takes the form:

$$\begin{aligned} y_j^* &= z\gamma_j + \eta_j,\ j = 1 \ldots M \\ y_1 &= x\beta_1 + u_1, \end{aligned} \tag{11}$$

where (η_j)'s are independent and identically type I extreme value distributed. Then the difference $\eta_j - \eta_j^{'}$ can be shown to be logistic distributed. u_1 does not have a parametrically specified distribution, but we know it has mean $E(u_1|x, z) = 0$ and variance $V(u_1|x, z) = \sigma^2$. j is a categorical variable that describes the choice of an agent among M alternatives based on utilities y_j^*. In my case, there are three alternatives, Moody's, S&P and both, so that $M = 3$ and $j = 1, 2, 3$. Without loss of generality, the outcome variable y_1 is observed if and only if category 1 is chosen, which happens when

$$y_1^* > \max_{j \neq 1}(y_j^*)$$

Define

$$\varepsilon_1 = \max_{j \neq 1}(y_j^*) - y_1$$

Choice y_1 is chosen when $\varepsilon_1 < 0$. As shown by McFadden (1973), the probability of choosing y_1 is

$$P(y_1|z) = P(\varepsilon_1 < 0|z) = \frac{\exp(z\gamma_1)}{\sum_j \exp(z\gamma_j)}$$

Based on this expression, consistent estimates of (γ_j)'s can be obtained by maximum likelihood method.

The problem is to estimate the parameter β_1 while taking into account that disturbance u_1 might not be independent of all (η_j)'s. It might introduce some correlation between the explanatory variable x and disturbance u_1 in the outcome equation, so that the OLS estimates of β_1 would be inconsistent.

Define $\Gamma = \{z\gamma_1, z\gamma_2, \cdots, z\gamma_M\}$. Generalizing the Heckman (1979) model, the bias correction term can be

$$E(u_1|\varepsilon_1 < 0, \Gamma) = \int\int_{-\infty}^{0} \frac{u_1 f(u_1, \varepsilon_1|\Gamma)}{P(\varepsilon_1 < 0|\Gamma)} d\varepsilon_1 du_1 = \lambda(\Gamma),$$

where $f(u_1, \varepsilon_1|\Gamma)$ is the conditional joint density of u_1 and ε_1. Therefore, consistent estimates of β_1 can be obtained by running OLS on regression:

$$y_1 = x_1\beta_1 + \lambda(\Gamma) + w_1 \tag{12}$$

where w_1 is a residual independent of regressors. The bias corrections among the four approaches differ in the assumptions over $\lambda(\Gamma)$. The assumptions between Lee and Dubin and McFadden's models are compared later.

- Lee's Model

  In Lee's model, there are two assumptions:

  A1. Correlation assumption:
  $Corr(u_1, \Phi^{-1}(P_1))$ does not depend on Γ.

  A2. Linearity assumption:
  $E(u_1|\varepsilon_1, \Gamma) = -\sigma\rho\Phi^{-1}(P_1)$

  Let the probability of selecting both Moody's and S&P be P_{ms}. The added bias correction terms in the second-stage equation are different across alternatives. Let the bias correction term added in the second-stage be m_i for alternative i. Then m_3 denotes the added regressor in the outcome equation of choosing both Moody's and S&P. It is equivalent to the inverse mills ratio in the Heckman's two-stage estimation of a bivariate selection model. Lee's method defines m_3 as

  $$m_3 = -\frac{\phi(\Phi^{-1}(P_{ms}))}{P_{ms}},$$

  where ϕ is the standard normal density, and Φ^{-1} is the inverse of cumulative normal distribution function. By running OLS on equation (13) consistent estimates of β_2 and σ_{12} can be obtained.

  $$y_{ms} = x\beta - \sigma_{12}\frac{\phi(\Phi^{-1}(P_{ms}))}{P_{ms}} + w_1, \tag{13}$$

where subscript i for issuer is omitted from each variable for succinctness. The correlation of the two errors can be estimated by $\rho = \frac{\sigma_{12}}{\sigma}$, where σ is the standard error of u_1. The variance of ε_{ij} in regression (2) can be estimated by the following equation:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{w}_1^2 + \sigma_{12}^2 V_3),$$

where $V_3 = m_3(\Phi^{-1}(P_{ms}) + m_3)$, and $\hat{w}_1$ is the OLS residual from the regression (13).

If $\rho \neq 0$, it means the two error terms are correlated, and sample selection correction is needed. The standard errors of estimates can be computed by bootstrap.

- Dubin and McFaddens Model
  The assumption of Dubin and McFadden's model is

  A3. DMF's linearity assumption:

$$E(u_1|\eta_1, \cdots, \eta_M) = \sigma\frac{\sqrt{6}}{\pi}\sum_{j=1,\cdots,M}\gamma_j(\eta_j - E(\eta_j)),$$

  where r_j is a correlation coefficient between u_1 and η_j. With the multinomial logit model (Dubin and McFadden (1984))

$$E(\eta_1 - E(\eta_1)|y_1^* > \max_{s\neq 1}(y_s^*), \Gamma) = -\ln(P_1)$$
$$E(\eta_j - E(\eta_j)|y_1^* > \max_{s\neq 1}(y_s^*), \Gamma) = \frac{P_j\ln(P_j)}{1-P_j}$$

  Let the added regressors in the second-stage equation be m_{1dmf_0}, m_{2dmf_0}, and m_{3dmf_0}.

$$m_{1dmf_0} = \frac{P_m\ln P_m}{1-P_m} + \ln P_{ms}$$
$$m_{2dmf_0} = \frac{P_s\ln P_s}{1-P_s} + \ln P_{ms}$$
$$m_{3dmf_0} = \ln(P_{ms})$$

The second-stage regression becomes

$$\begin{aligned} y_{ms} &= x\beta + \sigma\frac{\sqrt{6}}{\pi}(r_1 m_{1dmf_0} + r_2 m_{2dmf_0} - r_3 m_{3dmf_0}) + w_1 \\ &= x\beta + \sigma\frac{\sqrt{6}}{\pi}(r_1\frac{P_m \ln P_m}{1-P_m} + r_2\frac{P_s \ln P_s}{1-P_s} - r_3\ln(P_{ms})) + w_1 \end{aligned} \tag{14}$$

where subscript i for issuer is omitted from each variable for succinctness. The variance of ε_{ij} in regression (2) can be estimated by

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{w}_1^2 + \gamma_1^2 V_{1dmf_0} + \gamma_2^2 V_{2dmf_0}),$$

where

$$\begin{aligned} V_{1dmf_0} &= P_m(\frac{\ln P_m}{1-P_m})^2 \\ V_{2dmf_0} &= P_s(\frac{\ln P_s}{1-P_s})^2 \end{aligned}$$

and $\hat{w}_1$ is OLS residual from the regression (14).

The correlation between the errors of two stages is $\rho_i = \frac{\gamma_i \pi}{\sqrt{6}\sigma}$ for $i = m, s, ms$.

- Variant 1 of DMF model

  If we further assume that

  A4. $\sum_{j=1,\cdots,M}\gamma_j = 1$

  we will have

$$\begin{aligned} m_{1dmf_1} &= \frac{P_m \ln P_m}{1-P_m} + \ln(P_{ms}) \\ m_{2dmf_1} &= \frac{P_s \ln P_s}{1-P_s} + \ln(P_{ms}) \end{aligned}$$

  The second stage regression becomes

$$\begin{aligned} y_{ms} &= x\beta + \sigma\frac{\sqrt{6}}{\pi}(r_1 m_{1dmf_0} + r_2 m_{2dmf_0}) + w_1 \\ &= x\beta + \sigma\frac{\sqrt{6}}{\pi}(r_1(\frac{P_m \ln P_m}{1-P_m} + \ln(P_{ms}) + r_2(\frac{P_s \ln P_s}{1-P_s} + \ln(P_{ms})) + w_1 \end{aligned} \tag{15}$$

where subscript i for issuer is omitted from each variable for succinctness. The variance of ε_{ij} in regression (2) can be estimated by

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{w}_1^2 + \gamma_1^2 V_{1dmf_0} + \gamma_2^2 V_{2dmf_0}),$$

where

$$V_{1dmf_0} = P_m(\frac{\ln P_m}{1-P_m})^2$$
$$V_{2dmf_0} = P_s(\frac{\ln P_s}{1-P_s})^2$$

and $\hat{w}_1$ is the OLS residual from the regression (15).

The correlation between the errors of two stages is $\rho_i = \frac{\gamma_i \pi}{\sqrt{6}\sigma}$ for $i = m, s$, and $\rho_{ms} = -\rho_m - \rho_s$ for $i = ms$.

- Variant 2 of DMF model

  Define the standard normal variables $\eta_j^* = \Phi^{-1}(G(\eta_j)), j = 1, \cdots, M$, where (η_j)'s independently and identically follow Gumbel distribution, or called type I extreme value distribution. They have density function $g(\eta) = \exp(-\eta - \exp^{-\eta})$ and CDF $G(\eta) = \exp(-\exp^{-\eta})$. For every j, assume that the expected values of u_1 and η_j^* are linearly related. Let r_j^* be the correlation between u_1 and η_j^*, the expectation of u_1 can be expressed in the following linear combination:

  A5. Normalized DMF's linearity assumption

$$E(u_1|\eta_1, \cdots, \eta_M) = \sigma \sum_{j=1,\cdots,M} \gamma_j^* \eta_j^*$$

  and

$$m(P_j) = \int \Phi^{-1}(\nu - \ln(P_j))g(\nu)d\nu, \forall j$$

  Then performing some linear algebra yields

$$E(\eta_1^*|y_1^* > \max_{s\neq 1}(y_s^*), \Gamma) = m(P_1)$$
$$E(\eta_j^*|y_1^* > \max_{s\neq 1}(y_s^*), \Gamma) = \frac{m(P_j)P_j}{P_j - 1}, \forall j > 1$$

In my case,

$$
\begin{aligned}
m_{1dmf_2} &= m(P_m)\frac{P_m}{P_m - 1} \\
m_{2dmf_2} &= m(P_s)\frac{P_s}{P_s - 1} \\
m_{3dmf_2} &= m(P_{ms})
\end{aligned}
$$

The second stage regression becomes

$$
\begin{aligned}
y_{ms} &= x\beta + \sigma(\gamma_3^* m_{3dmf2} + \gamma_1^* m_{1dmf2} + \gamma_2^* m_{2dmf2}) \\
&= x\beta + \sigma(\gamma_3^* m(P_{ms} + \gamma_1^* m(P_m)\frac{P_m}{P_m - 1} + \gamma_2^* m(P_s)\frac{P_s}{P_s - 1})) \quad (16)
\end{aligned}
$$

where subscript i for issuer is omitted from each variable for succinctness. The variance of ε_{ij} in regression (2) can be estimated by

$$
\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{w}_1^2 + \gamma_1^2 V_{1dmf_2} + \gamma_2^2 V_{2dmf_2} + \gamma_3^2 V_{3dmf_2}),
$$

where

$$
\begin{aligned}
V_{1dmf_2} &= \frac{1 - P_m V_{1dmf_2}}{1 - P_m} - m_{1dmf_2}^2 \\
V_{2dmf_2} &= \frac{1 - P_s V_{2dmf_2}}{1 - P_2} - m_{2dmf_2}^2 \\
V_{3dmf_2} &= V_{3dmf_2} - m_{3dmf_2}^2
\end{aligned}
$$

and $\hat{w}_1$ is OLS residual from the regression (16).

The correlation between the errors of two stages is $\rho_i = \frac{\gamma_i}{\sigma}$ for $i = m, s, ms$.

In summary, the general procedure for estimating a multinomial selection model goes as follows. First, run a logit/probit regression of y_j on z and then obtain the probability of choosing y_1. Second, add the selection correction term into the second regression of y_1 on x using any of the four models above.

My peer effect model accounting for the selection bias cannot directly apply any of the above four approaches because the third alternative is to choose the combination of the first two products, Moody's and S&P. It causes the following complications in model design and estimation. First, an important variable used for identification is a rating agency's market share of an issuer in the last quarter. As an exclusion condition, this variable should be included in the first stage and excluded from the second stage. However, it is difficult to construct this variable in the utility equation of choosing both Moody's and S&P. Another complication is that the utility of choosing both Moody's and S&P is not simply the sum of utilities of choosing Moody's and S&P individually. A complementary or substitutionary effect need to be added. The last issue need to be taken into account is the possible correlation of the errors among the three choices and the correlation of the unobserved errors between the first and second stages.

5.2 Discrete Choice Model for Choosing Multiple Goods Simultaneously

Gentzkow (2007) developed a discrete-choice demand model that permits consumers to choose multiple goods simultaneously. Compare to large recent studies on discrete-choice models, his model enjoys the advantage of allowing for the complementary/substitutionary effect among products. He applied his model to study the impact of online newspaper on print newspaper. The choice set includes three goods: the Washington Post print edition, the post.com, and the Washington Times. Each consumer can choose any combination of the three goods. If a consumer chooses multiple goods, then the utility is the sum of the utility of each individual goods and a constant Γ which accounts for the complementary/substituitionary effect. The goods are complements if Γ is greater than zero, independent if Γ equals zero, and substitutes if Γ is less than zero. Based on his model, I set the utility of choosing both Moody's and S&P as the sum of the utility of choosing each of them separately and a complementary/substitution effect. This complementary/substitution effect is to be estimated.

Another feature of Gentzkow (2007) is his estimation method. He does not put any assumption the correlation matrix of the errors in the utility equations of different products. Instead, he allows the relationship between each pair of products to be freely estimated from the data. I cannot use his estimation method due to my data limitation. In the following part, I briefly introduce his estimation method, and then compare his method with mine.

Index days by t, consumers by $i = 1, \ldots, N$, goods by $j \in 1, \ldots, J$, and the set of possible bundles of these goods by $r \in 0, 1, \ldots, 2^J$. Assume the bundles are ordered so that $r = 0$ refers to the empty bundle and $r \in [1, J]$ refers to the singleton bundle consisting of only good $j = r$. The base utility to consumer i of consuming a single good j on day t is

$$\bar{u}_{ijt} = -\alpha p_j + \delta_j + x_i \beta_j + v_{ij} + \tau_{it}$$

where x_i is a vector of observable consumer characteristics, and v_{ij} and τ_{it} are unobservables. v_{ij} is assumed to have a J-dimensional multivariate normal distribution with a free covariance matrix.

The log-likelihood function is

$$L(x, q, \theta) = \sum_i \ln \int_{v_i} P_{q_i}(x_i, v; \theta) dF(v; \theta)$$

where $F(v; \theta)$ is the multivariate normal distribution of v conditional on parameters θ.

The straightforward way to estimate the log-likelihood function is to draw v from its distribution $F(v; \theta)$ for S times, and then compute the average of $P_{q_i}(x_i, v; \theta)$ over the S draws. The problem is that the covariance matrix of v is unknown.

Gentzkow (2007) deals with this problem in the following way. First, for each consumer i, he draws v_i from the i.i.d. standard normal distribution. v_i is a 1×3 vector, since the utility equation of each of three goods need a v_i. Second, he computes the probability of 7-day choice with the drawn v_i. Third, he draws a random variable u from the uniform distribution $U(0, 1)$. If u is less than the 7-day choice probability, the drawn v_i is accepted; otherwise, it is rejected. Given v_i and θ_0, he can estimate

$$\hat{P}_{q_i}(x_i; \theta_0) = \int_v P_{q_i}(x_i, v; \theta_0) dF(v; \theta_0)$$

by its approximation $\frac{1}{S} \sum_s P_{q_i}(x_i, v_i; \theta_0)$.

His estimation method is not applicable to my model since there is no observed data like 7-day choices to estimate the covariance matrix freely. Therefore, my covariance matrix of the choice errors is parameterized and estimated from the model. I used the GHK (Geweke-Hajivassiliou-Keane) simulator, named after Geweke (1989), Hajivassiliou and Mcfadden (1998) and Keane (1994), to approximate the integrals of choice probabilities. The regular GHK method is illustrated by Terracol (2002) as follows.

5.3 GHK Method

Consider a choice set y_1, y_2, y_3, the triprobit model supposes that

$$y_1 = \begin{cases} 1 & \text{if } X\beta + \varepsilon_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$
$$y_2 = \begin{cases} 1 & \text{if } Z\gamma + \varepsilon_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$
$$y_3 = \begin{cases} 1 & \text{if } W\theta + \varepsilon_3 > 0 \\ 0 & \text{otherwise} \end{cases}$$

where

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \sim N(0, \Sigma) \tag{17}$$

Suppose we want to evaluate $Pr(\varepsilon_1 < b_1, \varepsilon_2 < b_2, \varepsilon_3 < b_3)$, where $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are normal random variables with covariance matrix given in equation (17).

$$Pr(\varepsilon_1 < b_1, \varepsilon_2 < b_2, \varepsilon_3 < b_3) = Pr(\varepsilon_1 < b_1)Pr(\varepsilon_2 < b_2|\varepsilon_1 < b_1)Pr(\varepsilon_3 < b_3|\varepsilon_1 < b_1, \varepsilon_2 < b_2) \tag{18}$$

Instead of simulating $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ from the multivariate normal distribution, we could use Cholesky decomposition to transform to simulating random variables from an i.i.d. standard normal distribution.

Let L be the lower triangular Cholesky decomposition of Σ, such that $LL' = \Sigma$.

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

we have

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \tag{19}$$

where v_i are independent standard normal random variables. By equation (19) we have

$$\varepsilon_1 = l_{11}v_1$$
$$\varepsilon_2 = l_{21}v_1 + l_{22}v_2$$
$$\varepsilon_3 = l_{31}v_1 + l_{32}v_2 + l_{33}v_3$$

Thus,

$$
\begin{aligned}
Pr(\varepsilon_1 < b_1) &= Pr(v_1 < b_1/l_{11}) \\
Pr(\varepsilon_2 < b_2|\varepsilon_1 < b_1) &= Pr(v_2 < (b_2 - l_{21}v_1)/l_{22}|v_1 < b_1/l_{11}) \\
Pr(\varepsilon_3 < b_3|\varepsilon_1 < b_1, \varepsilon_2 < b_2) &= Pr(v_3 < (b_3 - l_{31}v_1 - l_{32}v_2)/l_{33}|v_1 < b_1/l_{11}, \\
&\qquad v_2 < (b_2 - l_{21}v_1)/l_{22})
\end{aligned}
$$

We draw a random variable v_1^* from a truncated standard normal density with upper truncation point of b_1/l_{11}, and v_2^* from a standard normal density with upper truncation point of $(b_2 - l_{21}v_1^*)/l_{22}$. Then equation (18) can be written as:

$$
Pr(v_1 < b_1/l_{11})Pr(v_2 < (b_2 - l_{21}v_1^*))Pr(v_3 < (b_3 - l_{31}v_1^* - l_{32}v_2^*)/l_{33}) \quad (20)
$$

The GHK simulator is the arithmetic mean of the probability given by equation (20) for D random draws of v_1^* and v_2^*:

$$
Pr_{GHK} = \frac{1}{D}\sum_{d=1}^{D} \Phi(b_1/l_{11})\Phi[(b_2 - l_{21}v_1^{*d})/l_{22}]\Phi[(b_3 - l_{31}v_1^{*d} - l_{32}v_2^{*d})/l_{33}]
$$

where v_1^{*d} and v_2^{*d} are the d-th draw of v_1 and v_2, and $\Phi(\cdot)$ is the CDF of normal distribution.

The next section is my model specification and estimation by GHK method.

Chapter 6

6 Model

6.1 Model Specification

Let i denote a single MBS bond, j denote a credit rating agency(CRA). For bond i at the beginning of each period t, CRA j can choose a rating R_{ijt}. All the possible ratings are from a choice set with 21 ratings, corresponding to rating AAA to D.

My model has two stages. In the first stage, an issuer selects which CRAs to rate the bond at issuance. In the second stage, the selected CRAs monitor the bond and update the rating when necessary till the bond's maturity. My interest is in the second stage - the peer effects on rating decisions after bond issuance. But we need to take into account the sample selection bias caused by the first stage selection of CRAs. There might be some unobservables which affect both the first-stage selection and the second-stage rating decisions. Therefore, we need to deal with the selection bias in the outcome equation.

Define the utility of an issuer for selecting CRA j to rate bond i as

$$U_{i,j} = \beta_{1i} X_i + \beta_{2j} F_j + \xi_j + v_{ij}, \tag{21}$$

where X_i is a vector of observed characteristics of bond i, including the original balance, the original support[15], the coupon rate, whether it is a fixed or floating rate bond, the number of tranches in a deal, and year and month effect; F_j is a vector of observed characteristics of CRA j, including CRAj's market share in the current quarter[16], the expected rating from CRA j based on bond's original support[17]; ξ_j is CRA j's unobserved characteristics; v_{ij} is unobserved variants in utility. I assume that CRA j's unobserved characteristics ξ_j not only affects the selection decision, but also the rating decision as

[15] Original support is measured in percentage

[16] Specifically, it is a CRA's market share of an issuer in the current quarter. A market is defined as the total number of bonds issued by a particular issuer in a given quarter. The market share of a rating agency in a market is then defined as the proportion of the total number of bonds it has rated in this market.

[17] First, find the quartiles of original support of all the bonds. In each quartile, calculate CRA j's average rating up to the previous quarter of bond i's issuance date.

well. Intuitively, ξ_j reflects the possibility that a CRA inflates ratings, which is known only by issuers who have done business with that CRA, but is not known by econometricians. If an issuer knows a CRA which might inflate ratings, the issuer will more likely select that CRA to rate his/her bonds, and will more likely obtain higher ratings later.

An issuer has three choices - select Moody, or S&P, or both to rate a bond. The utility of selecting CRA j to rate bond i is denoted as $U_{i,j}$. According to Gentzkow (2007)'s discrete choice model that allows consumers to choose multiple goods simultaneously, my utility functions take the following form which accounts the complementary/substitutionary effect when both Moody's and S&P are selected.

$$U_{i,A} = \beta_{1A} X_i + \beta_{2A} F_A + \xi_A + v_{i,A} \tag{22}$$

$$U_{i,B} = \beta_{1B} X_i + \beta_{2B} F_B + \xi_B + v_{i,B} \tag{23}$$

$$U_{i,AB} = (\beta_{1A} X_i + \beta_{2A} F_A + \xi_A + v_{i,A}) + (\beta_{1B} X_i + \beta_{2B} F_B + \xi_B + v_{i,B}) + \Gamma, \tag{24}$$

where Γ accounts for the complementary/substitutionary effect. Moody's and S&P are substitutes if $\Gamma < 0$, independent if $\Gamma = 0$, and complements if $\Gamma > 0$. I assume that $\xi_A \sim N(0, \sigma_{\xi_A})$, $\xi_B \sim N(0, \sigma_{\xi_B})$, and

$$(v_{iA}, v_{iB}) \sim N[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{AB} \\ \sigma_{AB} & 1 \end{bmatrix}] \tag{25}$$

In the second stage, I model the rating decisions as follows. Assume that there are only two firms in the market, Moody's and S&P.

$$(U_{i,A,t} | S_{i,A} = 1, S_{i,B} = 0) = \alpha_1 X_i + \eta_1 F_A + \xi_A + \epsilon_{i,A} \text{ – Moody is selected} \tag{26}$$

$$(U_{i,B,t} | S_{i,A} = 0, S_{i,B} = 1) = \alpha_2 X_i + \eta_2 F_B + \xi_B + \epsilon_{i,B} \text{ – S\&P is selected} \tag{27}$$

$$(U_{i,A,t} | S_{i,A} = 1, S_{i,B} = 1) = \gamma_1 |R_{i,A,t} - R_{i,B,t-1}| + \alpha_3 X_i + \eta_3 F_A + \xi_A + \epsilon^A_{i,AB} \text{ – Both are selected} \tag{28}$$

$$(U_{i,B,t} | S_{i,A} = 1, S_{i,B} = 1) = \gamma_2 |R_{i,B,t} - R_{i,A,t-1}| + \alpha_4 X_i + \eta_4 F_B + \xi_B + \epsilon^B_{i,AB} \text{ – Both are selected,} \tag{29}$$

where $\epsilon_{i,A}, \epsilon_{i,B}, \epsilon^A_{i,AB}, \epsilon^B_{i,AB}$ assume to be i.i.d extreme value distribution.

$R_{i,j,t}$ is the rating of bond i rated by CRA j at time t. $R_{i,j',t-1}$ is the rating of the other CRA on bond i at time $t-1$. Each rating agency can observe the other agency's rating in the last period when making its own decision.

I use maximum likelihood method to estimate the model. $P(S_{i,A} = 1, S_{i,B} = 0)$, $P(S_{i,A} = 0, S_{i,B} = 1)$, and $P(S_{i,A} = 1, S_{i,B} = 1)$ denote the probability that Moody's, S&P, and both are selected respectively.

For example, if Moody's is selected, we have the choice probability

$$\begin{aligned}
P(S_{i,A}=1,S_{i,B}=0) &= P(U_{iA}>U_{iB},U_{iA}>U_{iAB},U_{iA}>0|\ \max\{U_{iA},U_{iB},U_{iAB}\}>0) \\
&= \frac{P(U_{iA}>U_{iB},U_{iA}>U_{iAB},U_{iA}>0,\ \max\{U_{iA},U_{iB},U_{iAB}\}>0)}{P(\ \max\{U_{iA},U_{iB},U_{iAB}\}>0)} \\
&= \frac{P(U_{iA}>U_{iB},U_{iA}>U_{iAB},U_{iA}>0)}{P(\ \max\{U_{iA},U_{iB},U_{iAB}\}>0)} \\
&= \frac{P(U_{iA}>U_{iB},U_{iA}>U_{iAB},U_{iA}>0)}{1-P(U_{iA}<0,U_{iB}<0,U_{iAB}<0)}
\end{aligned}$$

$$\begin{aligned}
&P(U_{iA}>U_{iB},U_{iA}>U_{iAB},U_{iA}>0) \\
&\quad = P(v_{iB}-v_{iA}<-[(\beta_{1B}X_i+\beta_{2B}F_B+\xi_B)-(\beta_{1A}X_i+\beta_{2A}F_A+\xi_A)], \\
&\qquad v_{iB}<-(\beta_{1B}X_i+\beta_{2B}F_B+\xi_B)-\Gamma, v_{iA}>-(\beta_{1A}X_i+\beta_{2A}F_A+\xi_A)) \\
&\quad = \int_v I\{U_{iA}>U_{iB},U_{iA}>U_{iAB},U_{iA}>0\}dF(v;\theta)
\end{aligned}$$

where $F(v;\theta)$ is the multivariate normal distribution of v conditional on parameter θ. v_{iA} and v_{iB} are simulated from the joint normal distribution. The above choice probability is calculated via GHK method, due to Geweke et al. (1994), Hajivassiliou et al. (2010), and Keane (1994). The estimation process in detail is shown in Section 8 Estimation. In a similar way, I can estimate the probabilities $P(S_{i,A} = 0, S_{i,B} = 1)$ and $P(S_{i,A} = 1, S_{i,B} = 1)$.

The probability for CRA j to choose rating R_{ijt} at time t is

$$P(R_{i,A,t}, S_{i,A} = 1, S_{i,B} = 0) = P(R_{i,A,t}|S_{i,A} = 1, S_{i,B} = 0)P(S_{i,A} = 1, S_{i,B} = 0) \tag{30}$$

$$P(R_{i,B,t}, S_{i,A} = 0, S_{i,B} = 1) = P(R_{i,B,t}|S_{i,A} = 0, S_{i,B} = 1)P(S_{i,A} = 0, S_{i,B} = 1) \tag{31}$$

$$P(R_{i,A,t}, S_{i,A} = 1, S_{i,B} = 1) = P(R_{i,A,t}|S_{i,A} = 1, S_{i,B} = 1)P(S_{i,A} = 1, S_{i,B} = 1) \tag{32}$$

$$P(R_{i,B,t}, S_{i,A} = 1, S_{i,B} = 1) = P(R_{i,B,t}|S_{i,A} = 1, S_{i,B} = 1)P(S_{i,A} = 1, S_{i,B} = 1), \tag{33}$$

where $R_{i,j,t}$ is selected from the choice set with 21 ratings. Since ϵ are i.i.d extreme value distribution, the probability of agency j choosing each possible rating for a bond rated by two agencies can be written as:

$$P(Rating = R_{i,j,t}|S_{i,j} = 1, S_{i,j'} = 1) = \frac{\exp(\gamma_j|R_{i,j,t} - R_{i,j',t-1}| + \alpha_j X_i + \eta_j F_j + \xi_j)}{\sum_{k=1}^{21} \exp(\gamma_j|k - R_{i,j',t-1}| + \alpha_j X_i + \eta_j F_j + \xi_j)} \tag{34}$$

The maximum likelihood function is:

$$\begin{aligned} L(X, F; \boldsymbol{\theta}) &= \sum_i \sum_j \sum_t \ln P(R_{ijt}, R_{ij't}, S_{ij}, S_{ij'}) \qquad (35) \\ &= \sum_i \sum_j \sum_t \ln(P(R_{ijt}, R_{ij't}|S_{ij}, S_{ij'}) \times P(S_{ij}, S_{i,j'})) \\ &= \sum_i \ln P(S_{ij}, S_{ij'}) + \sum_t \sum_i \sum_j \ln P(R_{ijt}, R_{ij't}|S_{ij}, S_{ij'}), \qquad (36) \end{aligned}$$

where $\boldsymbol{\theta}$ is the parameter vector which includes all the parameters to be estimated in the model. In equation (36), the first part is the log-likelihood of first stage (i.e. selection stage) and the second part id the log-likelihood of the second stage (i.e. rating stage).

6.2 Model Identification

In this section, I discuss the identification of the model described by equation (21) to equation (29). To illustrate why the model can be identified, I start with explaining the intuition behind each variable in the model.

The utility equation (21) in the first stage has three independent variables. X_i is a vector of bond characteristics. Since the bond characteristics doesn't change with time, it does not have a subscript t. F_j include two variables, expected rating from CRA j based on bond's original support, and CRA j's market share of bond issuer i in the last quarter of the current business deal. The CRA's market share captures its relationship with the bond issuer. The more business they have done the better relationship they have, and the more likely the agency will inflate the rating later. Therefore, the relationship described by a CRA's market share of an issuer indicates the rating inflation the CRA will give to the issuer, if such inflation exists. The coefficient parameter β_{2j} denotes the amount of rating inflation that the bond issuer is certain about based on its experience with that agency. There is also an unobserved amount of rating inflation which is uncertain to the bond issuer and captured by ξ_j. This amount of rating inflation is random, and thus ξ_j is assumed to follow $N(0, \sigma_{\xi_j})$. For an issuer, all the bonds in the same deal issued at the same time will have the same random amount ξ_j. A CRA's market share of an issuer changes with time, so as their relationship. But at the time point that a bond issuer sponsors a deal, his/her relationship with the selected agency is fixed and the same for all the bonds in that deal. Thus, all the bonds in the same deal would have the same amount of unobserved rating inflation determined by the bond issuer's relationship with the rating agency.

Besides the agency's market share in last quarter, relationship can also be measured by alternative variables such as the number of business that has been done between an issuer and an agency, and an agency's average market share of a bond issuer in last three quarters. This relationship variable is included in both first and second stages, since relationship affects two stages' decision. If a bond issuer has close relationship with an agency, that agency will be more likely selected in the first stage, and then more likely inflate the rating in the second stage.

Both ξ_j and $v_{i,j}$ are unobserved variants on the utility of a bond issuer, but they affect the utility from different aspects. ξ_j captures the unobserved agency characteristics that affects rating inflation, while $v_{i,j}$ is the unobserved bond characteristics associated with the rating agency. Since both ξ_j and $v_{i,j}$ follow normal distribution, their summation also follows normal distribution. In the estimation of the choice probability of this probit model, the error terms will disappear in the integration.

In the second stage, the utility of choosing a rating grade depends on almost the same set of variables as in the selection stage - bond characteristics X_i, agency's characteristics F_j, and the random amount of rating inflation ξ_j. The interaction term of tranche type and delinquency rate is added into bond characteristics to control for the quality change of bonds, and the expected rating from an agency is excluded from agency's characteristics as an exclusion condition. The distribution of ξ_j can be identified by the case that bonds in the same deal issued by the same issuer at the same time but have different ratings. Those bonds have the same bond characteristics and the same CRA's market share, the variation in ratings come from the random effect ξ_j. But ξ_j is unobserved, and assumed to follow normal distribution. Unlike the ξ_j in the first-stage's probit model which can be aggregated with the error term and integrated out in the choice probability, ξ_j in the second-stage's logit model cannot be aggregated with the extreme value distributed error term $\epsilon_{i,j}$. It stays in the choice probability of ratings equation (34) and the likelihood function. It is unclear if the second-stage multinomial logit model can be identified since this unobserved variable ξ_j in the likelihood function makes the multinomial logit model nonstandard. Hence, my estimation proceeds in two ways. The first way is to estimate the two-stage model as a whole, maximizing the likelihood function (35). The second way is to estimate the two stages of the model sequentially - firstly maximize the likelihood of the first stage (i.e. the first part in the likelihood function (36)), then apply ξ_j generated from the first stage to the second stage, and maximize the likelihood of the second stage (i.e. the second part in the likelihood function (36)). It turns out that the estimated peer effect parameters γ_m and γ_s have very close values in these two ways.

Another way to deal with the unobserved random effect ξ_j in the second stage's likelihood function is to remove it. Then the second stage becomes a standard logit model with the only unobserved error term $\epsilon_{i,j}$. The first and second stages' decisions were correlated by having ξ_j in both stages. But now ξ_j is removed from the second stage. I allow the first and second stages' decisions still correlated through CRA's market share. As mentioned above, a CRA's market share of a bond issuer captures their relationship and eventually affects the amount/probability of rating inflation. Thus the amount/probability of rating inflation correlates the two stages by affecting the probability of a CRA being selected in the first stage and what rating that CRA will give in the second stage. Denote the coefficients of CRA's market share in the second stage by ϕ_A for Moody's and ϕ_B for S&P. Assume that $\phi_A = d \cdot \eta_A$ and $\phi_B = d \cdot \eta_B$ as show in equation 40 and equation

41. Since the amount/probability of rating inflation inferred from CRA and issuer's relationship could affect the utility on selection and rating decisions by different measurement unit, a constant d is multiplied to the coefficients of CRA's market share in rating equations to differentiate the impacts of rating inflation on two stages. The modified model is specified as follows.

The first stage is:

$$U_{i,A} = \beta_A X_i + \eta_A F_A + \xi_A + v_{i,A} \tag{37}$$

$$U_{i,B} = \beta_B X_i + \eta_B F_B + \xi_B + v_{i,B} \tag{38}$$

$$U_{i,AB} = (\beta_A X_i + \eta_A F_A + \xi_A + v_{i,A}) + (\beta_B X_i + \eta_B F_B + \xi_B + v_{i,B}) + \Gamma, \tag{39}$$

The second stage is:

$$(U_{i,A,t}|S_{i,A} = 1, S_{i,B} = 1) = \gamma_1 |R_{i,A,t} - R_{i,B,t-1}| + \alpha_3 X_i + d \cdot \eta_A F_A + \epsilon^A_{i,AB} \quad \text{– Both are selected} \tag{40}$$

$$(U_{i,B,t}|S_{i,A} = 1, S_{i,B} = 1) = \gamma_2 |R_{i,B,t} - R_{i,A,t-1}| + \alpha_4 X_i + d \cdot \eta_B F_B + \epsilon^B_{i,AB} \quad \text{– Both are selected,} \tag{41}$$

d could be identified for the following reasons. Suppose that d has different values for F_A and F_B, so that in (40) it is $d_A F_A$ and in (41)it is $d_B F_B$. There are four dependent variables in the model - $U_{i,A}, U_{i,B}$ in the first stage, and $U_{i,A,t}, U_{i,B,t}$ in the second stage, which means there are four equations (37), (38), (40), and (41). There are four unknown variables d_A, d_B, η_A and η_B. Therefore, these four unknown variables can be solved from the four equations. To make the first and second stages correlated, I put a restriction that $d_A = d_B = d$, thus d is identifiable.

The advantage of the modified model described by equation (37) to equation (41) is that the unobserved random effect ξ_j does not appear in the likelihood of logit model. On one hand, it makes the second logit model surely identified. On the other hand, it keeps the correlation between the first and second stages, and gives the correlation a specific meaning that it comes from the relationship between bond issuer and rating agencies. The disadvantage of the modified model is that there is no unobserved correlation between the first and second stage. The estimation result is reported in Table 14.

Another identification issue arises in estimating the sample selection model. In theory, the independent variables in the selection equation (21) and the rating equation (29) can be the same, and the identification can be achieved through the non-linearity of the underlying distribution functions. However, this may lead to weak identification problem (Keane (1992)). Therefore, I impose an exclusion condition to identify the sample selection model that at least one explanatory variable in the selection equation is excluded from the rating equation. For example, the expected rating from CRAs affects a bond issuer's selection of CRAs, but does not affect the selected CRA's rating decision in the second stage. Hence, the expected rating can be used as an exclusion condition.

6.3 Model Estimation

The likelihood function can be written as

$$L(X, F; \theta) = \sum_i \ln P(S_{ij}, S_{ij'}) + \sum_t \sum_i \sum_j \ln P(R_{ijt}, R_{ij't} | S_{ij}, S_{ij'}) \quad (42)$$

The first part is the sum of log-likelihood of selected CRAs for all the bonds, and the second part is sum of the log-likelihood of observed ratings given by the selected CRAs. Technically, it is possible to either maximize the two parts of the likelihood function as a whole or separately. Therefore, I carried out the estimation in both ways and compared the results.

I start by illustrating how to estimate the log-likelihood of the first stage. In the first stage, GHK method[18] is applied to simulate the joint normal distribution. For each bond, I calculate the probability of choosing four alternatives respectively - Moody, S&P, both agencies, and an outside option. If bond i's choice is Moody for example, it means that Moody has the largest utility among all the alternatives.

The probability of choosing Moody is

$$\begin{aligned} P(decision = Moody) &= P(S_{i,A} = 1, S_{i,B} = 0) \\ &= P(U_{iA} > U_{iB}, U_{iA} > U_{iAB}, U_{iA} > 0 | \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0) \\ &= \frac{P(U_{iA} > U_{iB}, U_{iA} > U_{iAB}, U_{iA} > 0, \ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)}{P(\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)} \\ &= \frac{P(U_{iA} > U_{iB}, U_{iA} > U_{iAB}, U_{iA} > 0)}{P(\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)} \\ &= \frac{P(U_{iA} > U_{iB}, U_{iA} > U_{iAB}, U_{iA} > 0)}{1 - P(U_{iA} < 0, U_{iB} < 0, U_{iAB} < 0)}, \end{aligned}$$

[18] GHK method has been introduced in section 5.

where the numerator is

$$\begin{aligned}
&P(U_{iA} > U_{iB}, U_{iA} > U_{iAB}, U_{iA} > 0) \\
&\qquad = P(V_{iA} + v_{iA} > V_{iB} + v_{iB}, V_{iA} + v_{iA} > V_{iA} + v_{iA} + V_{iB} + v_{iB} + \Gamma, \\
&\qquad\quad V_{iA} + v_{iA} > 0) \\
&\qquad = P(v_{iB} - v_{iA} < V_{iA} - V_{iB}, v_{iB} < -V_{iB} - \Gamma, v_{iA} > -V_{iA})
\end{aligned}$$

The variance-covariance matrix of (v_{iA}, v_{iB}) is

$$\Omega = \begin{pmatrix} 1 & \sigma_{AB} \\ \sigma_{AB} & 1 \end{pmatrix}$$

By Cholesky decomposition, we have the lower triangular matrix L that takes the form:

$$L = \begin{pmatrix} c_{aa} & 0 \\ c_{ab} & c_{bb} \end{pmatrix}$$

Using this Cholesky factor, the unobserved v_{iA} and v_{iB} can be written as linear functions of i.i.d standard normal random variables:

$$\begin{aligned}
v_{iA} &= c_{aa}\eta_{i1} \\
v_{iB} &= c_{ab}\eta_{i1} + c_{bb}\eta_{i2}
\end{aligned}$$

where η_{i1} and η_{i2} are i.i.d. standard normal random variables. Replace v_{iA} and v_{iB} with η_{i1} and η_{i2} in the utility functions we have

$$\begin{aligned}
U_{iA} &= V_{iA} + c_{aa}\eta_{i1} \\
U_{iB} &= V_{iB} + c_{ab}\eta_{i1} + c_{bb}\eta_{i2} \\
U_{iAB} &= V_{iA} + c_{aa}\eta_{i1} + V_{iB} + c_{ab}\eta_{i1} + c_{bb}\eta_{i2} + \Gamma
\end{aligned}$$

and we have

$$\begin{aligned}
&P(U_{iA} > U_{iB}, U_{iA} > U_{iAB}, U_{iA} > 0) \\
&\qquad = P(c_{aa}\eta_{i1} > -V_{iA}, c_{ab}\eta_{i1} + c_{bb}\eta_{i2} < -V_{iB} - \Gamma, \\
&\qquad\quad c_{bb}\eta_{i2} + (c_{ab} - c_{aa})\eta_{i1} < V_{iA} - V_{iB}) \\
&\qquad = P(\eta_{i1} > \frac{-V_{iA}}{c_{aa}}, \eta_{i2} < \frac{-V_{iB} - \Gamma - c_{ab}\eta_{i1}}{c_{bb}}, \eta_{i2} < \frac{V_{iA} - V_{iB} - (c_{ab} - c_{aa})\eta_{i1}}{c_{bb}}) \\
&\qquad = (1 - \Phi(\frac{-V_{iA}}{c_{aa}})) \times \Phi(\min(\frac{-V_{iB} - \Gamma - c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{V_{iA} - V_{iB} - (c_{ab} - c_{aa})\eta_{i1}^*}{c_{bb}})
\end{aligned}$$

The above equation suggests drawing η_{i1} and η_{i2} recursively. First draw η_{i1}^* from $N(0,1)$ truncated at $(\frac{-V_{iA}}{c_{aa}}, \infty)$, then draw η_{i2}^* from $N(0,1)$ truncated at $(-\infty,\ \min(\frac{-V_{iB}-\Gamma-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{V_{iA}-V_{iB}-(c_{ab}-c_{aa})\eta_{i1}^*}{c_{bb}}))$.

In the same way, I can compute the denominator $1 - P(U_{iA} < 0, U_{iB} < 0, U_{iAB} < 0)$ as follows:

$$
\begin{aligned}
&P(U_{iA} < 0, U_{iB} < 0, U_{iAB} < 0)\\
&\qquad = P(V_{iA} + v_{iA} < 0, V_{iB} + v_{iB} < 0, V_{iA} + v_{iA} + V_{iB} + v_{iB} + \Gamma < 0)\\
&\qquad = P(v_{iA} < -V_{iA}, v_{iB} < -V_{iB}, v_{iA} + v_{iB} < -V_{iA} - V_{iB} - \Gamma)\\
&\qquad = P(c_{aa}\eta_{i1} < -V_{iA}, c_{ab}\eta_{i1} + c_{bb}\eta_{i2} < -V_{iB}, (c_{aa} + c_{ab})\eta_{i1} + c_{bb}\eta_{i2} < -V_{iA} - V_{iB} - \Gamma)\\
&\qquad = P(\eta_{i1} < \frac{-V_{iA}}{c_{aa}}, \eta_{i2} < \frac{-V_{iB} - c_{ab}\eta_{i1}}{c_{bb}}, \eta_{i2} < \frac{-V_{iA} - V_{iB} - \Gamma - (c_{aa} + c_{ab})\eta_{i1}}{c_{bb}})\\
&\qquad = \Phi(\frac{-V_{iA}}{c_{aa}}) \times \Phi(\min(\frac{-V_{iB} - c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{-V_{iA} - V_{iB} - \Gamma - (c_{aa} + c_{ab})\eta_{i1}^*}{c_{bb}}))
\end{aligned}
$$

I draw η_{i1} and η_{i2} recursively. First draw η_{i1}^* from $N(0,1)$ truncated at $(-\infty, \frac{-V_{iA}}{c_{aa}})$, then draw η_{i2}^* from $N(0,1)$ truncated at $(-\infty, \min(\frac{-V_{iB}-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{-V_{iA}-V_{iB}-\Gamma-(c_{aa}+c_{ab})\eta_{i1}^*}{c_{bb}}))$.

The choice probability of choosing Moody is

$$
P(S_{i,A} = 1, S_{i,B} = 0) = \frac{(1 - \Phi(\frac{-V_{iA}}{c_{aa}})) \times \Phi(\min(\frac{-V_{iB}-\Gamma-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{V_{iA}-V_{iB}-(c_{ab}-c_{aa})\eta_{i1}^*}{c_{bb}})}{1 - \Phi(\frac{-V_{iA}}{c_{aa}}) \times \Phi(\min(\frac{-V_{iB}-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{-V_{iA}-V_{iB}-\Gamma-(c_{aa}+c_{ab})\eta_{i1}^*}{c_{bb}}))} \tag{43}
$$

Similarly, the probability of choosing S&P is

$$
\begin{aligned}
P(decision = S\&P) &= P(S_{i,A} = 0, S_{i,B} = 1)\\
&= P(U_{iB} > U_{iA}, U_{iB} > U_{iAB}, U_{iB} > 0 |\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)\\
&= \frac{P(U_{iB} > U_{iA}, U_{iB} > U_{iAB}, U_{iB} > 0,\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)}{P(\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)}\\
&= \frac{P(U_{iB} > U_{iA}, U_{iB} > U_{iAB}, U_{iB} > 0)}{P(\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)}\\
&= \frac{P(U_{iB} > U_{iA}, U_{iB} > U_{iAB}, U_{iB} > 0)}{1 - P(U_{iA} < 0, U_{iB} < 0, U_{iAB} < 0)},
\end{aligned}
$$

where

$$
\begin{aligned}
&P(U_{iB} > U_{iA}, U_{iB} > U_{iAB}, U_{iB} > 0) \\
&\qquad = P(V_{iB} + v_{iB} > V_{iA} + v_{iA}, V_{iB} + v_{iB} > V_{iA} + v_{iA} + V_{iB} + v_{iB} + \Gamma, V_{iB} + v_{iB} > 0) \\
&\qquad = P(v_{iA} - v_{iB} < V_{iB} - V_{iA}, v_{iA} < -V_{iA} - \Gamma, v_{iB} > -V_{iB}) \\
&\qquad = P((c_{aa} - c_{ab})\eta_{i1} - c_{bb}\eta_{i2} < V_{iB} - V_{iA}, c_{aa}\eta_{i1} < -V_{iA} - \Gamma, c_{ab}\eta_{i1} + c_{bb}\eta_{i2} > -V_{iB}) \\
&\qquad = P(\eta_{i1} < \frac{-V_{iA} - \Gamma}{c_{aa}}, \eta_{i2} > \frac{-V_{iB} - c_{ab}\eta_{i1}}{c_{bb}}, \eta_{i2} > \frac{V_{iA} - V_{iB} + (c_{aa} - c_{ab})\eta_{i1}}{c_{bb}}) \\
&\qquad = \Phi(\frac{-V_{iA} - \Gamma}{c_{aa}}) \times (1 - \Phi(\max(\frac{-V_{iB} - c_{ab}\eta_{i1}^*}{c_{bb}}, \\
&\qquad\quad \frac{V_{iA} - V_{iB} + (c_{aa} - c_{ab})\eta_{i1}^*}{c_{bb}})))
\end{aligned}
$$

Also, I draw η_{i1} and η_{i2} recursively. First draw η_{i1}^* from $N(0,1)$ truncated at $(-\infty, \frac{-V_{iA}-\Gamma}{c_{aa}})$, then draw η_{i2}^* from $N(0,1)$ truncated at $(\max(\frac{-V_{iB}-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{V_{iA}-V_{iB}+(c_{aa}-c_{ab})\eta_{i1}^*}{c_{bb}}), +\infty)$.

Then the probability of choosing S&P is

$$
P(decision = S\&P) = \frac{\Phi(\frac{-V_{iA}-\Gamma}{c_{aa}}) \times (1 - \Phi(\max(\frac{-V_{iB}-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{V_{iA}-V_{iB}+(c_{aa}-c_{ab})\eta_{i1}^*}{c_{bb}})))}{1 - \Phi(\frac{-V_{iA}}{c_{aa}}) \times \Phi(\min(\frac{-V_{iB}-c_{ab}\eta_{i1}^*}{c_{bb}}, \frac{-V_{iA}-V_{iB}-\Gamma-(c_{aa}+c_{ab})\eta_{i1}^*}{c_{bb}}))} \tag{44}
$$

The probability of choosing both Moody and S&P is

$$
\begin{aligned}
P(decision = Both) &= P(S_{i,A} = 1, S_{i,B} = 1) \\
&= P(U_{iAB} > U_{iA}, U_{iAB} > U_{iB}, U_{iAB} > 0 | \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0) \\
&= \frac{P(U_{iAB} > U_{iA}, U_{iAB} > U_{iB}, U_{iAB} > 0, \ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)}{P(\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)} \\
&= \frac{P(U_{iAB} > U_{iA}, U_{iAB} > U_{iB}, U_{iAB} > 0)}{P(\ \max\{U_{iA}, U_{iB}, U_{iAB}\} > 0)} \\
&= \frac{P(U_{iAB} > U_{iA}, U_{iAB} > U_{iB}, U_{iAB} > 0)}{1 - P(U_{iA} < 0, U_{iB} < 0, U_{iAB} < 0)},
\end{aligned}
$$

where

$$
\begin{aligned}
&P(U_{iAB} > U_{iA}, U_{iAB} > U_{iB}, U_{iAB} > 0) \\
&\qquad = P(V_{iA} + v_{iA} + V_{iB} + v_{iB} + \Gamma > V_{iA} + v_{iA}, V_{iA} + v_{iA} + V_{iB} + v_{iB} + \Gamma > V_{iB} + v_{iB}, \\
&\qquad\quad V_{iA} + v_{iA} + V_{iB} + v_{iB} + \Gamma > 0) \\
&\qquad = P(v_{iB} > -V_{iB} - \Gamma, v_{iA} > -V_{iA} - \Gamma, v_{iA} + v_{iB} > -V_{iA} - V_{iB} - \Gamma) \\
&\qquad = P(c_{ab}\eta_{i1} + c_{bb}\eta_{i2} > -V_{iB} - \Gamma, c_{aa}\eta_{i1} > -V_{iA} - \Gamma, \\
&\qquad\quad c_{aa}\eta_{i1} + c_{ab}\eta_{i1} + c_{bb}\eta_{i2} > -V_{iA} - V_{iB} - \Gamma) \\
&\qquad = (1 - \Phi(\frac{-V_{iA} - \Gamma}{c_{aa}}))(1 - \Phi(\max(\frac{-V_{iB} - \Gamma - c_{ab}\eta^*_{i1}}{c_{bb}}, \\
&\qquad\quad \frac{-V_{iA} - V_{iB} - \Gamma - (c_{aa} + c_{ab})\eta^*_{i1}}{c_{bb}})))
\end{aligned}
$$

Also, I draw η_{i1} and η_{i2} recursively. First draw η^*_{i1} from $N(0,1)$ truncated at $(\frac{-V_{iA}-\Gamma}{c_{aa}}, \infty)$, then draw η^*_{i2} from $N(0,1)$ truncated at $(\max(\frac{-V_{iB}-\Gamma-c_{ab}\eta^*_{i1}}{c_{bb}}, \frac{-V_{iA}-V_{iB}-\Gamma-(c_{aa}+c_{ab})\eta^*_{i1}}{c_{bb}}), \infty)$.

Finally the probability of choosing both agencies is

$$
P(decision = Both) = \frac{(1 - \Phi(\frac{-V_{iA}-\Gamma}{c_{aa}}))(1 - \Phi(\max(\frac{-V_{iB}-\Gamma-c_{ab}\eta^*_{i1}}{c_{bb}}, \frac{-V_{iA}-V_{iB}-\Gamma-(c_{aa}+c_{ab})\eta^*_{i1}}{c_{bb}})))}{1 - \Phi(\frac{-V_{iA}}{c_{aa}}) \times \Phi(\min(\frac{-V_{iB}-c_{ab}\eta^*_{i1}}{c_{bb}}, \frac{-V_{iA}-V_{iB}-\Gamma-(c_{aa}+c_{ab})\eta^*_{i1}}{c_{bb}}))} \tag{45}
$$

Now I illustrate how to estimate the choice probability in the second stage. The second stage is a multinomial logit model. To ensure the model identification, the highest rating (rating grade = 1) is chosen as the base category and its coefficients of all the case-specific regressors are set to zero. The coefficients of other ratings are then interpreted with respect to rating 1.

For example, the probability of CRA j choosing rating $R_{i,j,t}$ on bond i at time t is

$$
P(Rating = R_{i,j,t} | S_{i,j} = 1, S_{i,j'} = 1) = \frac{\exp(\gamma_j |R_{i,j,t} - R_{i,j',t-1}| + \alpha_j X_i + \eta_j F_j + \xi_j)}{\sum_{k=1}^{21} \exp(\gamma_j |k - R_{i,j',t-1}| + \alpha_j X_i + \eta_j F_j + \xi_j)} \tag{46}
$$

The result of a multimonial logit model is equivalent to a series of pairwise logit models. Then the multinomial logit defined in equation (28) and (29) imply that

$$
\begin{aligned}
Pr(Rating = R_{i,j,t}|Rating = R_{i,j,t} \text{ or } 1) &= \frac{Pr(Rating = R_{i,j,t})}{Pr(Rating = 1) + Pr(Rating = R_{i,j,t})} \\
&= \frac{\exp(\gamma_j |R_{i,j,t} - R_{i,j',t-1}| + \alpha_j X_i + \eta_j F_j + \xi_j)}{\exp(\gamma_j |1 - R_{i,j',t-1}|) + \exp(\gamma_j |R_{i,j,t} - R_{i,j',t-1}| + \alpha X_i + \eta F_j + \xi_j)}
\end{aligned}
$$

using α_j and η_j are zero when rating $= 1$ and the cancelation of $\sum_{k=1}^{21} \exp(\gamma|k - R_{i,j',t-1}| + \alpha X_i + \eta F_j + \xi_j)$ in the numerator and denominator. So a positive coefficient from multinomial logit means that as the regressor increases, it is more likely to choose alternative rating $R_{i,j,t}$ than rating 1.

The model is estimated in four ways:

- Estimate the two stages as a whole, using a full set of 48 parameters. The likelihood function (42) is maximized as a whole. The estimation result is shown in Table 10.

- Estimate the two stages as a whole, but let Moody's and S&P share the same set of parameters in the rating stage. It combines equation (28) and (29) into one equation below and reduces the parameter set into 37 parameters.

$$
(U_{i,j,t}|S_{i,A} = 1, S_{i,B} = 1) = \gamma |R_{i,j,t} - R_{i,j',t-1}| + \alpha X_i + \eta F_j + \xi_j + \epsilon^{j}_{i,AB} \tag{47}
$$

  The above equation assumes that Moody's and S&P have the same impact on each other's rating decision by γ, and they are equally affected by bond characteristics X_i and agency characteristics F_j. The estimation result is shown in Table 11.

- Estimate the two stages sequentially, using a full set of 48 parameters. First estimate the first stage by maximizing the first part of the likelihood function, i.e. $\sum_i \ln P(S_{ij}, S_{ij'})$ in equation 36, and then use the parameters σ_{ξ_A} and σ_{ξ_B} estimated from the first stage to maximize the second part of the likelihood function 36, i.e. $\sum_t \sum_i \sum_j \ln P(R_{ijt}, R_{ij't}|S_{ij}, S_{ij'})$. The estimation result is shown in Table 12.

- Estimate the two stages sequentially, using a reduced set of 37 parameters. That is to use equation (47) in the second stage. The estimation result is shown in Table 13.

It is necessary to set up bounds for parameters in the optimization process. First, the variance σ_{ξ_A} and σ_{ξ_B} should be greater than zero. Second, the element c_{ab} in the lower triangular matrix of Cholesky decomposition should be no greater than 1, so that $c_{bb} = \sqrt{1 - c_{ab}^2}$ can be a real number.

Denote the full parameter set by $\boldsymbol{\theta}$. The distribution of maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ can be approximated by a multivariate normal distribution with mean equals to the true parameters $\boldsymbol{\theta}_0$ and covariance matrix equals to $\{\boldsymbol{I}(\boldsymbol{\theta}_0)\}^{-1}$, where

$$\begin{aligned} \boldsymbol{I}(\boldsymbol{\theta}_0) &= -E_0[\frac{\partial^2 \ln f(X_i;\hat{\theta})}{\partial\theta_0\partial\theta_0^{'}}] \\ &= -\frac{1}{N}\sum_{i=1}^{N}\frac{\partial^2 \ln f(X_i;\hat{\theta})}{\partial\theta_0\partial\theta_0^{'}} \end{aligned} \tag{48}$$

Matlab optimization function *fmincon* stores the second derivative of log-likelihood function in variable *hessian*. Therefore, the variance of $\hat{\boldsymbol{\theta}}$ is the reciprocal of *hessian*. It is simple to implement this numerical method to compute standard error. But one disadvantage is that the hessian matrix could be close to singular and cannot have inverse.

Chapter 7

7 Results

The full parameter set include 48 parameters $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4, \boldsymbol{\eta}_3, \boldsymbol{\eta}_4, \Gamma, \sigma_{\xi_A}, \sigma_{\xi_B}, \gamma_1, \gamma_2, \sigma_{AB}\}$. The parameters in bold are vectors. In the first stage, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are coefficients of bond characteristics and agency characteristics respectively; Γ is the complement/substitute effect; σ_{ξ_A} and σ_{ξ_B} are the variance of the distribution of ξ_A and ξ_B; σ_{AB} is the covariance between ϵ_A and ϵ_B; In the second stage, γ_1 and γ_2 are the peer effects; $\boldsymbol{\alpha}_3$ and $\boldsymbol{\alpha}_4$ are the coefficients of bond characteristics and $\boldsymbol{\eta}_3$ and $\boldsymbol{\eta}_4$ are the coefficients of agency characteristics.

7.1 Results by Four Ways of Estimation

Table 10 shows the results of estimating the two stages as a whole by maximizing the sum of likelihood functions of both selection and rating equations. Column (1) has self-specified starting points and parameter bounds $[-10, 10]$. Column (2) has starting points randomly generated from a uniform distribution and bounds $[-20, 20]$. Column (3) and (4) have randomly generated starting points and no bounds. Column (4) also includes the interaction term of tranche type and 90-day plus delinquency rate. The peer effect is statistically significant for Moody and S&P around -0.946 and -0.963 respectively without delinquency rate in the rating equation. The negative value of peer effect means that the difference with competitor's rating decreases an agency's utility. One unit of rating difference with its competitor decreases Moody's and S&P's utility by 94% and 96% respectively. S&P is 1.7% more negatively affected by its rating difference with Moody's than Moody's being affected by its rating difference with S&P without considering the delinquency rate. When delinquency rate is included in the rating equation, the peer effect decreases from -0.946 to -0.929 on Moody's, and decreases from -0.963 to -0.948 on S&P. It means that the rating changes are partially explained by the change of delinquency rate on the underlying assets additional to the peer effect. Without considering the quality change of the underlying asset will overestimate the peer effect on agency's decisions. Adding the delinquency rate into the model also enlarges the difference in peer effects between Moody's and S&P by 1.9%. The positive interaction term of tranche type and 90-day plus delinquency rate indicates that delinquency rates increase the probability of choosing current rating instead of choosing rating one (i.e. best rating).

In the rating equation, the original balance and original support for both Moody and S&P are negative. It means that if original balance or support increases, an agency is less likely to choose current rating choice, but more likely to choose rating one (i.e. best rating). It makes sense given that if the original balance or support increases, a bond's credit risk will decrease, and thus it will more likely obtain a better rating. The impact of number of tranches is not clear, since it is negative in Moody's equation but positive in S&P's equation. The average market share has positive effect on both Moody and S&P's utility. It is significant for S&P in four columns, and significant for Moody only in Column (3). It indicates that if a CRA's market share of an issuer decreases in the last quarter before the current business deal, the CRA will tend to issue a generous rating to the issuer's bonds in order to gain more business. On the other hand, if the CRA's market share increases in the last quarter before the current business deal, the CRA will have less incentive to favor the bond issuer by inflating the rating. Note that including the interaction term have modest impact on the coefficient parameters of original balance and support and the average market share in the last quarter for both Moody's and S&P. The number of tranches in a deal in Moody's equation becomes significant after including the interaction term in the model.

In the selection equation part of Table 10, the coefficients of Moody's original balance and support are positive. It suggests that increasing original balance and support increases the probability of choosing Moody. The effect of S&P's original balance is positive while the effect of original support is negative. The values of S&P's original balance and support also vary across columns. The number of tranches in the deal has positive effect on both Moody's and S&P's utility, indicating the increases of number of tranches will increases the probability of selecting the two agencies. The expected rating decreases the utility as expected. The higher expected rating the worse rating is, and thus decreases the probability of choosing that agency. The expected rating is significant at 95% confidence level for both Moody's and S&P when delinquency rate is added A CRA's average market share of a bond issuer in the last quarter increases the probability of the CRA being selected by the bond issuer. The complementary/substitution effect is significantly positive, showing that choosing two rating agencies has complementary effect.

This complementary effect makes the utility of choosing two agencies greater than the sum of utility of choosing each individual agency. It also explains why as many as 88% of bonds in our data used for model estimation are rated by both agencies. The correlation of error terms σ_{AB} is positive, but non-significant. The unobserved characteristics ξ_A and ξ_B which influence both selection and rating stages have zero mean and close to zero variance. Therefore, the impact from the unobserved characteristics which correlates the two stages might be negligible.

Table 11 shows the results of estimating the two stages as a whole with a reduced set of parameters. Reduced parameters means that the same set of parameters is used for both Moody's and S&P's rating equations in the second stage. Column (1) reports the estimation result with self-picked starting points and parameter bounds $[-10, 10]$. Column (2) reports the estimation result with randomly generated starting points and bounds $[-20, 20]$. Column (3) and column (4) have randomly generated starting points and no parameter bounds. Column (4) also includes the interaction of tranche type and 90-day plus delinquency rate. The peer effect is statistically significant at -0.953 without considering the delinquency rate, while it is significant at -0.937 considering the delinquency rate. It means that the absolute value of one unit increase in rating difference decreases the utility by 93.7% and 95.3 with and without the delinquency rate respectively. The exclusion of delinquency rate will overestimate the peer effect by 1.6%. Compared to the base choice rating grade one, an increase in original balance and support will increase the probability of choosing rating one (i.e. the best rating). Number of tranches in the deal has non-significant positive effect on utility. The average market share in last quarter of the business deal has significantly negative effect on utility. The increase in average market share in last quarter will decrease the probability of choosing current rating, and increase the probability of choosing rating one. This impact is opposite from the result of estimating the two stages as a whole with full parameters in Table 10. The interaction term of tranche type and delinquency rate are positive for all tranches, indicating an increase in the delinquency rate will increase the probability of having a worse rating.

In the selection stage of Table 11, the original balance, original support and the number of tranches in the deal have significant positive effect on the utility of choosing Moody's, but have no significant effect on the utility of choosing S&P. The impact of expected rating is -0.076 on Moody's and -0.178 on S&P without the delinquency rate. When the delinquency rate is considered, the impact of expected rating stays almost the same for Moody's, but decreases to -0.353 for S&P. It confirms that the higher expected rating (i.e. worse rating) the lower the utility is. The average market share in last quarter is around 1.8 and 2.448 for Moody's with and without the delinquency rate, and is around 4.4 and 2.078 for S&P with and without the delinquency rate. It indicates that an issuer tend to choose the agency which has taken higher of its market share. Conventionally when an agency has done more recent business with an issuer than other agencies, the building-up relationship with the issuer makes that agency more likely to be selected. When delinquency rate is considered, the impact of market share increases on the utility of choosing Moody's but decreases on the utility of choosing S&P. Choosing both agencies have complementary effect by the amount around 1.3. The unobserved characteristics which affects both selection and rating stages might not exist, since σ_{ξ_A} and σ_{ξ_B} have zero mean and close to zero variance. The correlation between the error terms of v_{iA} and v_{iB} is around 0.01, but is not significant.

Table 12 shows the results of estimating the two stages sequentially with a full set of parameters. This method starts with estimating the first stage, and then applies the random effect ξ_A and ξ_B generated from the first stage to the second stage. Column (1) uses $[-20, 20]$ as the parameter's bounds for the first stage estimation, and uses no bounds in second stage. The starting point is self-specified. Column (2) also uses self-specified starting points but no bounds in the estimation. Column (3) uses randomly generated starting points and no bounds. Column (4) includes the interaction term of tranche type and 90-day plus delinquency rate with randomly generated starting points and no bounds. The negative maximum likelihood is around 87000 in the first three columns and 77000 in the last column with the delinquency rate included in the rating equation. The peer effect is -0.946 for Moody's and -0.960 for S&P without considering the delinquency rate. When delinquency rate is included, the peer effect decreases to -0.927 for Moody's and to -0.948 for S&P. It is the same as previous estimation result that S&P is more affected by Moody's decision than Moody's being affected by S&P's decision. Such difference in peer effect between the two agencies is 1.4% without considering the delinquency rate. Including the delinquency rate enlarges the difference of peer effects between Moody's and S&P to 2.1%. In the rating equation, the original balance, original support, and number of tranches in the deal for Moody are all significant, with values around -0.14, -13.04, and -0.05 respectively in the first three columns. They are consistent with the estimation results in Column (2) and Column (3) from Table 10 which estimates the model as a whole with a full set of parameters. The original balance and original support for S&P are also negatively significant, meaning increasing the original balance and support will increase the quality of bonds and thus decrease the probability of choosing current rating and moving to a better rating. The number of tranches in the deal for S&P is positive but not significant. The interaction term of tranche type and 90-day plus delinquency rate is positive for junior and mezzanine tranches but negative for senior tranch. It is different from the results of other three ways of estimation in which the sign of interaction terms are all positive. It suggests that the impact of delinquency rate on tranches depends on the tranche type. An increase in delinquency rate means deterioration of junior and mezzanine tranches and then worse rating. But the senior tranche type might dominate the interaction term, so that an increase in delinquency rate does not affect the rating much but the senior tranche type increase the chance of having a better rating.

In the selection equation of Table 12, Moody's original balance, support and number of tranches are all significant around the values 0.03, 9.20 and 0.05 respectively either with or without the delinquency rate in the model. Since the two stages are estimated separately and the delinquency rate is only included in the second stage, the factors in selection equation is not supposed to be affected by including the delinquency rate. The Moody's expected rating are all negative, while S&P's expected rating are negative in the last two columns but insignificantly positive in the first two columns. It is consistent with the fact that a higher rating grade (i.e. worse rating) gives an issuer lower utility. The average market share in last quarter are all significantly positive, suggesting that expanding the market share increases an agency's chance of being selected. S&P's original balance and support are positive but not significant. The number of tranches in the deal significantly and positively affects S&P's chance of being selected in the last three columns. The complementary effect from choosing two agencies is significantly positive with values vary between 1 and 1.5 across the four columns. The positive value is in line with previous result that choosing two agencies have greater utility than choosing each individual agency. The correlation between error terms v_{iA} and v_{iB} ranges from 0.3 to 0.8, and is only significant when delinquency rate is included. The variance of the unobserved random effect on two stages is close to zero for Moody's, and is not significant for both Moody and S&P. It suggests that the correlations between the two stages decisions might not exist.

Table 13 gives the results of estimating two stages sequentially with a reduced set of parameters. As in Table 11, Moody's and S&P share the same set of parameters in the rating stage. Column (1) uses $[-10, 10]$ as the parameter bounds in the first stage estimation, and use no bounds in the second stage. The starting points are self-specified according to parameter's economic intuition. Column (2) uses randomly generated starting points, and parameter bounds $[-20, 20]$ in the first stage but no bounds in the second stage. Column (3) uses randomly generated starting points and no bounds in both stages. Column (4) includes the interaction term of tranche type and 90-day plus delinquency rate with randomly generated starting points and no bounds. The total negative maximum likelihood is around 86000 in the first three columns without delinquency rate, and is 75917 in the fourth column including the delinquency rate in the model. The peer effect is significantly negative at -0.953 without delinquency rate. It means the absolute value of one unit difference with competitor's rating decreases the agency's utility by 95.3%. When the delinquency rate is included in the model, the peer effect decreases to -0.937. It confirms with the previous results that delinquency rate explains the rating changes in addition to the peer effect. Without considering the impact of delinquency rate will cause the peer effect upward biased by 1.6%. In the rating equation, the original balance, the original support, and the number of tranches in a deal are significantly negative around -0.11, -9.50, and -0.01 respectively either with or without delinquency rate in the model. Delinquency rate does not affect the impact of those factors on Moody and S&P's rating decisions. The average market share in last quarter is significantly negative around -2.8. It suggests that an agency tends to issue a favorable rating to issuers of which the agency has a large market share in recent business. The original balance and support, the number of tranches in a deal, and an agency's average market share in the last quarter are all close to the results in Table 11 estimating the two stages as a whole with reduced parameters. The interaction terms of tranche type and delinquency rate are all positive, indicating that an increase in the delinquency rate increase the probability of having a worse rating.

In the selection stage in Table 13, Moody's original balance, the original support, and the number of tranches in a deal are significantly positive around 0.03, 9.20, and 0.05 respectively either with or without delinquency rate in the model. Since the two stages are estimated separately and the delinquency rate is only in the second stage, the factors in the first stage will not be affected by including delinquency rate. S&P's original balance and number of tranches in the deal are positive but not significant. S&P's original support are significantly negative around -2.01 in Column (1) and Column (3). Moody's average market share in the last quarter is significantly positive around 1.8. S&P's average market share is positive in all columns, but only significant in Column (1) and (3). The values in Column (1) and (3) are around 4.38, which is very close to the results in Table 11 estimating the two stages as a whole with reduced parameters. A positive coefficient of average market share means that an increase in average market share in last quarter increases an agency's probability of being selected. Moody's expected rating is significantly negative around -0.07, while S&P's expected rating are not significant in all columns. The complementary effect of choosing two agencies is significant and the value varies between 1 and 1.5 in the four columns. The correlation between the error terms are not significant. The variance of unobserved random effect ξ_A and ξ_B included in both stages are close to zero and not significant. It shows that there is no significant correlation between the two stages' decisions. Also, the unobserved characteristics ξ_A and ξ_B might have very small impact on both stages due to their zero mean and insignificant and close-to-zero variance.

Table 14 shows the estimation results for the modified model from Equation (37) to Equation (41). Column (1) uses estimation bound $[-20, 20]$ and the self-specified starting points. Column (2) uses random starting points, and no bound. The results are close to Table 10 and Table 12estimating two stages both sequentially or as a whole with a full set of parameters. The peer effect is significant for both Moody and S&P at -0.946 and -0.963, respectively. Moody's original balance, original support, and the number of tranches in a deal in rating equation are significantly negative at -0.142, -12, and -0.052, respectively. Moody's average market share in last quarter is significantly positive at 1.88. S&P's original balance and original support in rating equation are significantly negative at -0.090 and -4.38, respectively. The number of tranches in a deal has non-significant coefficient 0.07. The average market share in last quarter are also positive, and significant in estimation without bounds. In the selection equation, Moody's original balance, support and number of tranches in a deal are all significantly positive as results in previous tables either estimating the two stages as a whole or sequentially, with a full or reduced set of parameters. The same is true for S&P's original balance, support and number of tranches in a deal except that they are not significant. The complementary effect is positive. The common factors ξ_A and ξ_B taken into account in both stages do not have significant impact. The correlation between error terms v_{iA} and v_{iB} is not significant either. The newly added parameter d, which is used to differentiate the utility measure in selection and rating stages, is only significant under the estimation with bound $[-20, 20]$. Given that its value is positive and less than one, the rating inflation in rating equation has less impact on utility than in selection equation.

The above four ways of model estimation suggest that the estimates for the peer effects are consistent. Without including the interaction term of tranche type and delinquency rate in the rating equation, Moody's and S&P's peer effect is around -0.946 and -0.963 respectively in the estimation of a full set of parameters; in the estimation of a reduced set of parameters, the peer effect is around -0.953. When the interaction term of tranche type and delinquency rate is included in the rating equation, Moody's and S&P's peer effect reduces to -0.929 and -0.948 respectively in the estimation of a full set of parameters; in the estimation of a reduced set of parameters the peer effect reduces to -0.937. In summary, delinquency rate explains the rating changes in addition to peer effect. Without including the delinquency rate in the model overestimates the peer effect by around 2%. S&P is 1.7% more affected by Moody's rating decision than Moody's being affected by S&P's decision without considering the delinquency rate. When delinquency rate is included in the rating equation, this differentials between Moody's and S&P's peer effect on each other increases to 2.2%. In the rating decision, the coefficients of original balance and support, and the number of tranches in a deal are consistent for both Moody and S&P. The three variables are all significantly negative as expected. Since the increase in original balance, or original support, or the number of tranches will improve the bond quality, and thus increase the probability of obtaining a better rating. In the selection equation, the coefficient of original balance, support, number of tranches in a deal and average market share are consistent in the estimation of the reduced set of parameters for both Moody and S&P, but differs in the estimation of the full set of parameters for S&P. The average market share in the last quarter is significantly positive as expected. It means that an issuer is more likely to choose an agency which has its higher market share. The expected rating has negative effect on selection decision, since a higher expected rating grade (i.e. worse rating) gives an issuer lower utility. The value of likelihood function varies if the estimation bounds on parameters change. When delinquency rate is not considered in the model, the value of log-likelihood is around 86000 with a full set of parameters and around 85000 with a reduced set of parameter. When the delinquency rate is considered, the log-likelihood reduces to around 75000 either with a set of full or reduced parameters. Year and tranche dummies are included in the model but not reported. Their values vary with the bounds on parameters and the way of estimation.

7.2 Model Identification and Estimation Assessment

As stated above, the model is estimated in four ways - estimate the two stage as a whole with a full set of parameters and with a reduced set of parameters, and estimate the two stage sequentially with a full set of parameters and with a reduced set of parameters. In each way of estimation, I start with a self-picked starting points and upper and lower bounds in the optimization process, and then relax those constraints. Therefore, there are four columns reported in each way of estimation: (1) Self-picked starting points and bounds (2) Self-picked starting points and no bounds (3) Random starting points and no bounds (4) Random starting points and no bounds with interaction term of tranche type and delinquency rate. It shows that the rating stage has reasonably consistent estimates especially for the parameter of peer's rating, while some parameter estimates in the selection stage are affected by the way of estimation, starting points, and whether the bounds are added. When different starting points or bounds are chosen, the original support, expected rating, and average market share in last quarter have varied estimates in S&P's selection equations. Nevertheless, their signs are consistent so that we could know the direction in which these variables affect the selection decision. It suggests that there is difficulty for those variables to be identifiable. By comparing the parameter estimates across the four ways of estimation, the parameter estimates are close except the original support and number of tranches in a deal in S&P's rating equation have small difference between estimating two stages as a whole and estimating two stages sequentially.

To ensure the identification of the rating equation, the unobserved variable ξ_A and ξ_B are removed and thus make the rating equation a standard logit model with only the error terms unknown. Table 14 shows the estimation results of the modified model described by equation (37) to equation (41). In the modified model, the unobserved random effects ξ_A and ξ_B which are designed to correlate the two stages' decision are removed from the second stage. ξ_A and ξ_B in the first stage remain to capture the unobserved agency characteristics which affects issuer's selection decision. Removing the unobserved variables from the second stage is to ensure the identification of the logit model. The correlation between the first and second stage are instead captured by an agency's market share. An agency has a large market share of a bond issuer implies their close relationship, which makes the agency more likely being selected in the first stage and more likely to inflate the rating in the second stage. Since the market share might affects the selection decision and rating decision by different extent, the coefficient of market share in the second stage is the multiplication of the coefficient of market share in the first stage by a constant d. The estimated peer effects are the same as the previous model. The peer effects on Moody's and S&P are -0.946 and -0.963 without delinquency rates, and are -0.929 and -0.948 with delinquency rate. It means that the previous model can be identified with the unobserved variables in the second stage logit model. The constant variable d which measures the different extents that market share affect the selection and rating decision is only significant when estimated at chosen starting points and without delinquency rate. The complementary effect of choosing two agencies is still positive but not significant in the modified model.

7.3 Estimation Results of Model without Selection

In Table 15 and Table 16, the model is estimated without the selection of rating agencies by bond issuers. In other words, ξ_A and ξ_B are set to zero, and only the second-stage of the model described by equation (28) and equation (29) is estimated. To be comparable with results of model with selection, both a full set of parameters and a reduced set of parameters are estimated and reported in Table 15 and Table 16 respectively. In these two tables, Column (1) uses self-picked starting points and bounds $[-30, 30]$ on parameters. Column (2) uses starting points randomly generated from a uniform distribution, and bounds $[-30, 30]$. Column (3) uses randomly generated starting points and no bounds. Column (4) includes the interaction term of tranche type and 90-day plus delinquency rate, randomly generated starting points and no bounds.

In Table 15 with a full set of parameters, the peer effect without selection stage is almost the same as the one with selection stage either with or without the interaction term of tranche type and delinquency rate. It is consistent with our findings that the selection decision and the rating decision are not strongly correlated given that ξ_A and ξ_B have mean zero and variance close to zero from our estimation. Besides the peer effect, most of other parameters are all close to the results in Table 10 and Table 12 which estimate the model with selection stage. In Table 16 with a reduced set of parameters, the finding is the same as in Table 15 with a full set of parameters that the peer effect without selection stage is almost the same as the one with selection stage in Table 11 and Table 13. It means that the two-stage's decisions also have little correlation when the model is described by a reduced set of parameters.

7.4 Marginal Effect of Peer's Rating

This section analyzes the marginal effect of peer's rating. Table 17 and Table 18 present the probabilities that an agency downgrades/upgrades the rating given that its peer downgrades/upgrades the rating by one notch. Table 17 uses the parameter estimates in Column (4) of Table 10 which estimates the two stages as a whole. Table 18 uses the parameter estimates in Column (4) of Table 12 which estimates the two stages separately. The results based on the two ways of estimation are close.

Here are some noteworthy findings in Table 17 and Table 18. First of all, the magnitude of peer effects for different ratings is considerably non-linear. For AAA rating (i.e. rating grade = 1), the peer effect is small. When an agency observes that its peer's rating downgrades by one notch from AAA, there is only 1% $\sim$ 2% probability for that agency to downgrade its rating. However, for bonds with AA rating, the peer effects are much larger. The probability for an agency to downgrade its rating from AA increases to 66% $\sim$ 86% if its peer downgrades its rating by one notch. The magnitude of peer effect declines for A- rating. The probability of downgrading rating from A- given peer's downgrade by one notch declines to 44% for Moody's and 80.51% for S&P. From A- to BB+ rating, the downgrade probability stays around 45% for Moody's and 75% for S&P if the peer's rating were one notch lower. After B rating, the peer effect measured by downgrade probabilities goes down below 10% again. In other words, peer effects are close to zero for AAA bonds, then increase significantly for medium-rating bonds such as AA, and then decline again for lower-rating bonds such as A.

The second finding is that the upgrade probabilities is much lower than downgrade probability. For AA bonds, Moody's and S&P's downgrade probabilities around 66% and 87% had peer's rating been one notch lower are much higher than their upgrade probabilities around 11% and 34% had peer's rating been one notch higher. From A+ to BB bonds, Moody's and S&P's

downgrade probabilities are on average around 45% and 70% respectively had peer's rating been one notch lower, as compared to their upgrade probabilities on average around 7% and 20% had peer's rating been one notch higher. It suggests that peer effect is stronger on downgrading a bond than upgrading a bond.

Third, peer effect on S&P is more significant than on Moody's. Have observed the peer's downgrading/upgrading its bond by one notch, S&P's downgrade/upgrade probability is higher than Moody's on all the rating grades except those low-rating grades (below BB) with scarce data. S&P's downgrade and upgrade probabilities are greater than Moody's by approximately 30% and 13% respectively from A+ to BB bonds.

Lastly, downgrading probability at investment grade is not lower than around investment grade. By regulations certain institutional investors such as pension funds and insurance companies cannot hold bonds lower than investment grade. If bonds in their portfolios are downgraded below investment grade, the composition of the portfolio has to be adjusted by selling those non-investment grade bonds. With the concern of the loss and transaction cost caused by such adjustment, CRAs might be cautious about downgrading bonds below investment grade, and try to maintain the rating stability. Hence, the probability of downgrading bonds at investment grade is likely to be lower than the probability of downgrading bonds around investment grades. Table 17 and Table 18 disprove this argument, and show evidence that the downgrade probability at investment grade (rating grade = 10) is not lower than the downgrade probability at rating grade 9 or 11. It suggests that the inconvenience brought to investors might not affects CRA's decision on downgrading bonds below investment grade.

The rating change probabilities given peer's rating change by one notch are also compared between including and excluding the delinquency rate. Though the magnitude of downgrade/upgrade probabilities are not exactly the same at each rating, the trend is the same: peer effects are small for AAA bonds, increases sharply for medium-rating bonds and then declines for lower-rating bonds.

Chapter 8

8 Conclusion

Credit rating inflation in sub-prime mortgage backed securities has been accused of exacerbating the financial crisis in 2008 and defrauding investors by offering overly favorable ratings to MBS. There are two underlying sources for the rating inflation - credit rating shopping by bond issuers and the peer effect among rating agencies. My paper applies a two-stage model to estimate the peer effect among CRAs while taking into account the selection process to avoid selection bias. It shows robust evidence that Moody and S&P's rating decision are affected by each other. When delinquency rate is not included in the rating equation, the peer effects on Moody's and S&P's rating decision is around -0.946 and -0.963 respectively. S&P is more affected by Moody's decision than Moody's being affect by S&P by 1.7%. When underlying asset's delinquency rate is considered in the rating stage, the peer effects on Moody's and S&P reduce to -0.927 and -0.948 respectively. But the difference in their peer effect increases to 2.1%. If Moody's and S&P use the same set of parameters in the rating equation, the peer effect is -0.953 and -0.937 with and without the delinquency rate respectively. It indicates that the change in delinquency rate of underlying assets partially explains the rating changes made by CRAs. Without considering the change in underlying assets' quality will overestimate the peer effects.

In the rating stage, the original balance and support are significantly negative. The increases of original balance and original support will increase the chance of choosing rating grade one (i.e. best rating), and decrease the chance of choosing current rating. A CRA's market share in last quarter before the current business deal is significantly positive in the estimation of using a full set of parameters. It means that the decrease in market share in last quarter will decrease the chance of current rating and thus increase the chance of best rating. It is consistent with our expectation, since the lower market share an agency has the more likely the agency wants to favor the bond issuer and gain business by inflating the rating. The impact of number of tranches in a deal is inclusive since it is significantly negative for Moody's rating but not significant for S&P's rating. The interaction term of tranche type and delinquency rate is positive for junior, mezzanine and senior tranches, indicating that the increase in the delinquency rate means the decease of the bond quality and thus increase the probability of obtaining a worse rating.

In the selection stage, the increase in original balance will increase the probability of choosing both agencies while the increase in original support only increase the probability of choosing Moody's but not S&P. The higher expected rating (worse rating) decreases issuer's utility, and thus the issuer will be less likely to choose that agency. The average market share in last quarter increases the utility of choosing both agencies. The multinomial selection model has significant complementary effect of choosing two agencies. The correlation between the error terms v_{iA} and v_{iB} are not significant. The unobserved agency characteristics ξ_A and ξ_B which affect both selection and rating decisions has mean zero and variance close to zero. It implies that there might be little unobserved agency characteristics that taken into account in both stages. The number of tranches in a deal has positive effect on selection decision and increase an issuer's utility.

If the model is estimated without the selection stage, the peer effect is almost the same as with the selection stage. It is in line with our estimation result that there might be no common unobserved agency characteristics that affects two stages, since ξ_A and ξ_B has zero mean and close-to-zero variance. The analysis on marginal effect of peer's rating has the following findings. Firstly, peer effects are small for AAA bonds, then increase significantly for medium-rating bonds (AA), and then decline again for lower-rating bonds (A). Secondly, Moody's and S&P's upgrade probabilities is much lower than their downgrade probability at all ratings. Thirdly, the peer effects on S&P is more significant than on Moody's. Lastly, the downgrading probability at investment grade is not lower than that around investment grade. In other words, whether a bond is at the investment grade does not affect an agency's downgrading decision.

There are at least two noteworthy caveats of my studies. First, I only look at the peer effect between Moody and S&P. The study can be extend to include four rating agencies, Moody, S&P, Fitch, and DBRS. For example, how one rating agency's rating decision can be affected by the other three? The peer effect might be different on a big-size agency from on a small-size agency. Second, it would be interesting to study the strategic interaction among rating agencies. Instead of looking at how one agency's rating in the last period affects the other's rating, I can study the static game between these two agencies, and solve for an equilibrium of their rating decisions. In this case, dealing with multiple equilibria would be a challenge.

References

Alessi, C., 2013: The credit rating controversy. campaign 2012. Council on Foreign Relations. Retrieved 29 May 2013.

Altman, E., 2001: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, **23 (4)**, pp. 589–609.

Ammermueller, A. and J. Pischke, 2009: Peer effects in european primary school: Evidence from pirls. *Journal of Labor Economics*, **27**, pp.315–348.

Ashcraft, A. and T. Schuermann, 2008: Understanding the securitization of subprime mortgage credit. Mimeo, Federal Reserve Bank of New York.

Atkins, P. S., 2008: Speech by sec commissioner: Remarks to the institute of international bankers, `http://www.sec.gov/news/speech/2008/spch030308psa.htm`, retrieved November 17, 2011.

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov, 2013: Estimating static models of strategic interaction, `http://www.nber.org/papers/w12013`.

Bajari, P. and J. Krainer, 2004: An empritical model of stock analysts' recommendations: Market fundamentals, conflicts of interest, and peer effects. *National Bureau of Economic Research*.

Bar-Isaac, H., 2003: Imperfect competition and reputational commitment. *Economics Letters*, **89 (2)**, pp. 167–173.

Becker, B. and T. Milbourn, 2011: How did increased competition affect credit ratings? *The Journal of Financial Economics*, **101**, pp. 493–514.

Benmelech, E. and J. Dlugosz, 2010: The credit rating crisis, national Bureau of Economic Research.

Bolton, P., F. Xavier, and S. Joel, 2012: The credit rating game. *The Journal of Finance*, **67**, pp. 85–112.

Brock, W. A. and S. N. Durlauf, 2007: Identification of binary choice models with social interactions. *Journal of Econometrics*, **140**, pp. 52–75.

Cameron, A. and P. K. Trivedi, 2005: *Microeconmetrics:Methods and Applications*. Cambridge.

Cantor, R. and M. Christopher, 2006: Analyzing the trade-off between ratings accuracy and stability. Special Comment on Moody's Investors Service.

Cantor, R. and F. Packer, 1997: Difference of opinion and selection bias in the credit rating agency. *Journal of Banking and Finance*, **21**, pp. 1395–1417.

Carey, M. and M. Hrycay, 2001: Parameterizing cedit risk models with rating data. *Journal of Banking and Finance*, **25**, pp.197–270.

Chomsisengphet, S. and A. Pennington-Cross, 2006: The evolution of the subprime mortgage market. *Federal Reserve Bank of St. Louis Review*, pp. 31–56.

Chu, C. and R. Fackler, 2013: Rating agency behavior in structured finance: Bid shading and peer effects, federal Reserve Board of Governors.

Dahl, G. B., 2002: Mobility and the returns to education: testing a roy model with multiple markets. *Econometrica*, **70 (6)**, pp. 2367–2420.

Dubin, J. A. and D. L. McFadden, 1984: An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, **52**, pp.345–362.

Ederington, L., 1985: Classification models and bond ratings. *The Financial Review*, **20**, pp. 237–261.

Edward, A. and R. Herbert, 2006: A point-in-time perspective on through-the-cycle ratings. *Financial Analyst Journal*.

Faltin-Traeger, O., 2009: Picking the right rating agency: Issuer choice in the abs market, columbia Busines School.

Followill, R. and T. Martell, 1997: Bond review and rating change announcements: An examination of informational value and market efficiency. *The Journal of Economics and Finance*, **21 (2)**, pp. 75–82.

Franois, B., F. Martin, and G. Marc, 2004: Selection bias corrections based on the multinomial logit model: Monte-carlo comparisons, DELTA Working Papers.

Gan, Y. and C. Mayer, 2007: Agency conflicts, asset substitution, and securitization, columbia Business School and NBER, working paper, June 11.

Gentzkow, M., 2007: Valuing new goods in a model with complementarity: Online newspaper. *American Economic Review*, **97 (3)**, pp. 713–744.

Geweke, J., 1989: Bayesian inference in econometric models using monte carlo integration. *Econometrica*, **57 (6)**, pp. 1317–1339.

Geweke, J., M. P. Keane, and D. Runkle, 1994: Alternative computational approaches to inference in the multinomial probit model. *The Review of Economics and Statistics*, **76 (4)**, pp.609–632.

Gonzales, F., F. Haas, R. Johannes, M. Persson, L. Toledo, R. Violi, C. Zins, and M. Wieland, 2004: Market dynamics associated with credit ratings: A literature review. *Banque de France Financial Stability Review*, **4**, pp. 53–76.

Gunter, L., 2004: An anatomy of rating through-the-cycle. *Journal of Banking and Finance*, **28**, pp.695–720.

Hajivassiliou, V., D. McFadden, and P. Ruud, 2010: Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results. *Journal of Business and Economic Statistics, American Statistical Association, American Statistical Association*, **28 (4)**, pp.469–482.

Hajivassiliou, V. A. and D. L. Mcfadden, 1998: The method of simulated scores for the estimation of ldv models. *Econometrica*, **66 (4)**, pp. 863–896.

He, J., Q. Jun, and E. Philip, 2003: Credit ratings and evolution of the mortgage-backed securities market. *American Economic Review: Papers and Proceedings*, **101**, pp. 131–135.

Heckman, J. J., 1979: Sample selection bias as a specification error. *Econometrica*, **47 (1)**, pp. 153–161.

Horner, J., 2002: Reputation and competition. *American Economic Review*, **92 (3)**, pp. 644–663.

John, K., K. Michael, and L. Schumacher, 2013: Rating through-the-cycle: What does the concept imply for rating stability and accuracy?, working paper, International Monetary Fund, Mar. 2013.

Jorion, P. and G. Zhang, 2007: Information effects of bond rating changes: The role of the rating prior to the announcement. *The Journal of Fixed Income*, **16 (4)**, pp. 45–59.

Kamstra, M., P. Kennedy, and T.-K. Suan, 2001: Combining bond rating forecasts using logit. *The Financial Review*, **37**, pp. 75–96.

Keane, M., 1992: A note on identification in the multinomial probit model. *Journal of Business and Economic Statistics*, **10**, pp. 193–200.

Keane, M. P., 1994: A computationally practical simulation estimator for panel data. *Econometrica*, **62 (1)**, pp.95–116.

Klein, B. and K. B. Leffler, 1981: The role of market forces in assuring contractual performance. *The Journal of Political Economy*, **89 (4)**, pp. 615–641.

Lee, L., 1983: Generalized econometric models with selectivity. *Econometrica*, **51**, pp.507–512.

Li, S., Y. Liu, and K. Deininger, 2013: How important are endogenous peer effects in group lending? estimating a static game of incomplete information. *Journal of Applied Econometrics*, **28 (5)**, pp. 864–882.

Mahlmann, T., 2011: Is there a relationship benefit in credit ratings? *Review of Finance*, **15**, pp. 475–510.

Manski, C., 2000: Economic analysis of social interactions. *The Journal of Economic Perspectives*, **14**, pp. 115–136.

McFadden, D., 1973: Conditional logit analysis of qualitative choice behavior. *in P. Zarembka(ed.)Froniters in Econometrics*, pp. 105–142.

Nelson, C., D. Pragyan, and Z. Liu, 2012: Credit rating and competition, working paper, presented at the 2011 AFA Annual Meeting in Denver, US.

Stefan, H., 2013: Credit rating dynamics and competition, `http://ssrn.com/abstract = 1961510`.

Sun, Q., S. Li, and W. Tan, 2013: An empirical analysis of credit ratings game in subprime mortgage-backed security market, working Papers.

Tan, W. and F. Wang, 2008: Joint ratings and credit rating quality, working paper.

Terracol, A., 2002: Triprobit and the ghk simulator: a short note, mimeo.

White, L. J., 2010: The credit rating agencies. *The Journal of Economic Perspectives*, **24 (2)**, pp. 211–226.

Xia, H., 2012: Can competition improve the information quality of credit ratings?

Zhou, C., 2001: Credit rating and corporate default. *Journal of Fixed Income*, **3 (11)**, pp. 30–40.

Table 1: Rating Grade

Numeric Grades	Moody's	S&P	Fitch	DBRS
1	Aaa	AAA	AAA	AAA
2	Aa1	AA+	AA+	AAH
3	Aa2	AA	AA	AA
4	Aa3	AA-	AA-	AAL
5	A1	A+	A+	AH
6	A2	A	A	A
7	A3	A-	A-	AL
8	Baa1	BBB+	BBB+	BBBH
9	Baa2	BBB	BBB	BBB
10	Baa3	BBB-	BBB-	BBBL
11	Ba1	BB+	BB+	BBH
12	Ba2	BB	BB	BB
13	Ba3	BB-	BB-	BBL
14	B1	B+	B+	BH
15	B2	B	B	B
16	B3	B-	B-	BL
17	Caa1	CCC+		
18	Caa2	CCC	CCC	
19	Caa3			
20	Ca	CC	CC	
21	C	C	C	C
22		D	D	

Table 2: Examples of bond observation

Variable Description	Bond 1	Bond 2
Cusip	1266715F9	004421MN0
Intex Deal Name	CWHE0404	ACE05HE2
Issuer	Countrywide ABS	ACE Securities Corp
Vintage	2004	2005
Closing Date	3/31/2004	3/29/2005
Tranche Name	2A	B1
Tranche Type	Senior Floater	Mezzanine Floater
Bond Original Balance($ millions)	340.0	16.5
Bond Original Support (%)	19.05	1.2
Coupon Rate (%)	5.52	8.755
Number of Tranches in the Deal	13	16
Moody's Original Rating	Aaa	Ba3
S&P Original Rating	AAA	BB+
Fitch Original Rating		BB+
DBRS Original Rating		
Moody's Rating at 1st Rating Change		Caa1
Moody's 1st Rating Change Date		5/16/2007
S&P Rating at 1st Rating Change		B
S&P 1st Rating Change Date		10/15/2007
Fitch Rating at 1st Rating Change		BB
Fitch 1st Rating Date		5/10/2007
Fitch Rating at 2nd Rating Change		C
Fitch 2nd Rating Date		12/10/2007

Table 3: Number of Bonds Rated by Multiple Rating Agencies

Rating Agencies	Number of Bonds	Percentage (%)
Moody's only	27	0.16
S&P only	619	3.56
Fitch only	98	0.56
DBRS only	13	0.07
Subtotal	757	4.35
Moody's and S&P	6599	38.00
Moody's and Fitch	228	1.31
Moody's and DBRS	9	0.05
S&P and Fitch	835	4.81
S&P and DBRS	63	0.36
Fitch and DBRS	23	0.13
Subtotal	7757	44.66
Moody's and S&P and Fitch	6793	39.11
Moody's and S&P and DBRS	542	3.12
Moody's and Fitch and DBRS	62	0.36
S&P and Fitch and DBRS	82	0.47
Subtotal	7479	43.06
Moody's and S&P and Fitch and DBRS	1374	7.91
Total	17367	100

Table 4: Rating Changes of Bonds Rated by Two Agencies

Bonds Rated by Two Agencies	Changed by 1st Only	Changed by 2nd Only	Changed by Both	1st Agency Changes First	2nd Agency Changes First	Change at the Same Time
Moody's and S&P	876	199	2740	1622	1099	19
Moody's and Fitch	31	19	99	23	76	0
Moody's and DBRS	1	0	8	7	1	0
S&P and Fitch	29	146	342	139	203	0
S&P and DBRS	1	10	33	13	20	0
Fitch and DBRS	2	0	13	5	8	0

Notes:

- This Table only included bonds rated by exactly two rating agencies. Bonds rated by three or more agencies are excluded.
- Insured Bonds are excluded.
- Take the first row "Moody's and S&P" for example. The first column "Changed by 1st Only" lists the number of bonds which have their ratings changed by Moody's only. The second column "Changed by 2nd Only" lists the number of bonds which have their ratings changed by S&P only. "Changed by Both" lists the number of bonds which have their ratings changed by both Moody's and S&P. "1st Agency Changes First" lists the number of bonds which have first rating change made by Moody's. "2nd Agency Changes First" lists the number of bonds which have first rating change made by S&P. "Change at the Same Time" lists the number of bonds which have first rating change made by Moody's and S&P at the same time.

Table 5: Rating Changes of Bonds Rated by Three Agencies

Bonds Rated by 3 Agencies	Changed by 1st Only	Changed by 2nd Only	Changed by 3rd Only	Changed by 1st+2nd	Changed by 2nd+3rd	Changed by 1st+3rd	Changed by All	1st Changes First	2nd Changes First	3rd Changes First	Change at Same Time
Moody's S&P Fitch	269	70	357	58	168	456	2348	648	332	1359	0
Moody's S&P DBRS	50	10	35	50	1	30	224	107	44	73	0
Moody's Fitch DBRS	4	1	0	2	5	7	34	10	14	10	0
S&P Fitch DBRS	0	5	1	3	9	2	42	7	14	21	0

Notes:

- This Table only included bonds rated by exactly three rating agencies. A bond rated by four agencies has been excluded.
- Insured Bonds are excluded.
- Take the first row "Moody's, S&P and Fitch" for example. The first column "Changed by 1st Only" lists the number of bonds which have their ratings changed by Moody's only. The second column "Changed by 2nd Only" lists the number of bonds which have their ratings changed by S&P only. 'Changed by 3rd Only" lists the number of bonds which have their ratings changed by Fitch only.
- The fourth column "Changed by 1st+2nd" lists the number of bonds which have their ratings changed by both Moody's and S&P. "Changed by 2nd+3rd" lists the number of bonds which have their ratings changed by both S&P and Fitch. "Changed by 1st+3rd" lists the number of bonds which have their ratings changed by both Moody's and Fitch. "Changed by All" lists the number of bonds which have their ratings changed by Moody's, S&P and Fitch.
- The eighth column "1st Changes First" lists the number of bonds which have first rating change made by Moody's. "2nd Changes First" lists the number of bonds which have first rating change made by S&P. "3rd Changes First" lists the number of bonds which have first rating change made by Fitch. "Changes at the Same Time" lists the number of bonds which have first rating change made by Moody's, S&P, and Fitch at the same time.

Table 6: Rating Changes of Bonds Rated by Four Agencies

Bonds Rated by 4 Agencies	Changed by Moody's Only	Changed by S&P Only	Changed by Fitch Only	Changed by DBRS Only	Changed by All	Moody's Changes First	S&P Changes First	Fitch Changes First	DBRS Changes First	Changes at Same Time
Moody's S&P Fitch DBRS	33	10	34	26	644	181	60	293	110	0

Notes: Insured Bonds are excluded.

Table 7: Frequency of Rating Changes

Number of Rating Changes	Moody's		S&P		Fitch		DBRS	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
0	7402	47.35	9340	55.24	4468	47.06	921	42.48
1	3434	21.96	4247	25.12	3620	38.13	496	22.88
2	3002	19.20	2714	16.05	1333	14.04	369	17.02
3	1730	11.07	566	3.35	72	0.76	261	12.04
4	59	0.38	40	0.24	1	0.01	107	4.94
5	7	0.04			1	0.01	14	0.65
Total	15634	100	16907	100	9495	100	2168	100

Notes: Insured Bonds are excluded.

Table 8: Market Coverage

Vintage	Original Balance($mm)					Market Coverage by Weighted Original Balance (%)			
	Moody's	S&P	Fitch	DBRS	Total	Moody's	S&P	Fitch	DBRS
2004	325776	351228	209617	20892	354168	37	41	20	2
2005	418538	440879	281219	51634	454150	36	39	22	3
2006	440077	440822	212676	61006	446130	41	41	15	4
Total	1184391	1232184	703512	133532	1254448	38	40	19	3
Vintage	Number of Bonds Rated					Market Coverage by Number of Bonds (%)			
	Moody's	S&P	Fitch	DBRS	Total	Moody's	S&P	Fitch	DBRS
2004	3850	4312	2486	237	4368	35	43	20	2
2005	5385	6056	3764	840	6329	34	41	21	4
2006	6778	6510	3048	988	6778	40	41	15	4
Total	16013	16878	9298	2065	17475	37	42	18	3

Notes:

- The weighted original balance of bond k is calculated as the original balance of bond k divided by the number of ratings on the bond.
- The market coverage of CRA i in vintage j by weighted original balance is calculated as the sum of the weighted original balance of bonds rated by CRA i in vintage j divided by the total original balance of all bonds in vintage j.
- Let an indicator equals one if bond k is rated by CRA i, and zero otherwise. The indicator can be used to calculate the total number of bonds rated by CRA i adjusted for the number of ratings on each bond. The adjustment is made by dividing the indicator by the number of ratings on the bond.
- The market coverage of CRA i in vintage j by number of bonds is calculated as the sum of the adjusted indicators if bond k is rated by CRA i in vintage j divided by the total number of bonds in vintage j.

Table 9: Variable Description and Summary Statistics

Variable	Description		Number of Observations	Mean	Median	Standard Deviation	Min	Max
Bond-level characteristics								
Initial Rating	Initial ratings for sub-prime mortgage tranches. 1 denotes the best rating, 21 denotes the worst rating	Moody's	5338	4.73	4	3.56	1	20
		S&P	5887	4.59	4	3.46	1	14
Original Balance	Balance of the bond when issued		5914	57.55	18.03	104.46	0	509.47
Original Support	The credit support for the bond when issued. It is measured as the percentage of the face value of the bond.		5914	11.60	10.05	7.87	0	66.5
Number of Tranches	Number of tranches that structured in the deal		5914	15.59	15	5.10	2	34
Fixed Rate	Equals one if the tranche is a fixed rated tranche; equals 0 if otherwise		5914	0.11	0	0.31	0	1
Coupon Rate	The yield of the tranche paid on its issue date		5914	6.19	5.82	0.94	2.92	12.45
CRA-level characteristics								
Average Market Share	CRA's average market share in the last three months	Moody's	40998	0.91	0.93	0.11	0	1
		S&P	40998	0.98	1	0.06	0.63	1
Expected Rating	CRA's average rating up to previous quarter within each quartile of original support	Moody's	40998	4.75	5.42	3.36	1.06	9.43
		S&P	40998	4.38	4.95	3.05	1.06	8.83

Table 10: Estimate Two Stages as a Whole - with Full Parameters

	(1)	(2)	(3)	(4)
Rating Equation				
Moody's				
Peer Ratings	-0.946**	-0.946**	-0.946**	-0.929**
	(0.009)	(0.009)	(0.011)	(0.007)
Original Balance	-0.148**	-0.143**	-0.143	-0.144**
	(0.029)	(0.037)	(0.120)	(0.014)
Original Support	-12.968	-13.046**	-13.046**	-17.267**
	(1.989)	(1.003)	(1.594)	(1.087)
Number of Tranches in the Deal	-0.050	-0.052	-0.052	-0.015**
	(0.049)	(0.063)	(0.078)	(0.008)
Average Market Share in Last 3 Months	0.515	0.733	0.738**	0.062
	(0.933)	(1.001)	(0.354)	(0.056)
Junior × 90-day plus Delinquency Rate				0.109**
				(0.020)
Mezzanine × 90-day plus Delinquency Rate				0.005
				(0.012)
Senior × 90-day plus Delinquency Rate				0.016
				(0.013)
S&P				
Peer Ratings	-0.963**	-0.963**	-0.963**	-0.948**
	(0.009)	(0.009)	(0.009)	(0.007)
Original Balance	-0.090**	-0.090**	-0.090**	-0.088**
	(0.055)	(0.039)	(0.052)	(0.016)
Original Support	-4.387**	-4.390**	-4.390**	-3.892**
	(1.427)	(1.003)	(0.614)	(0.437)
Number of Tranches in the Deal	0.007	0.007	0.007	0.017
	(0.059)	(0.063)	(0.055)	(0.012)
Average Market Share in Last 3 Months	1.574**	1.572*	1.574**	1.490**
	(0.605)	(1.001)	(0.532)	(0.126)
Junior × 90-day plus Delinquency Rate				0.011
				(0.080)
Mezzanine × 90-day plus Delinquency Rate				0.028**
				(0.008)
Senior × 90-day plus Delinquency Rate				0.045**
				(0.012)
Selection Equation				
Moody's				
Original Balance	0.026	0.026	0.026	0.028**
	(0.045)	(0.046)	(0.225)	(0.008)
Original Support	9.013**	9.156**	8.957**	10.448**
	(1.050)	(1.079)	(3.804)	(0.076)
Number of Tranches in the Deal	0.056**	0.049	0.049	0.058
	(0.000)	(0.039)	(0.175)	(0.009)
Expected Rating	-0.076	-0.064	-0.065	-0.077**
	(0.083)	(0.092)	(0.326)	(0.012)
Average Market Share in Last 3 Months	1.809**	1.884*	1.871	2.446**
	(0.076)	(1.302)	(1.534)	(0.106)
σ_{ξ_A}	0.001	0.000	0.000	0.001
	(0.003)	(1.000)	(1.000)	(0.994)
S&P's				
Original Balance	0.049**	1.982**	0.891**	0.045**
	(0.001)	(0.080)	(0.202)	(0.012)
Original Support	-2.048**	-4.038**	-0.367	-4.408**
	(0.000)	(1.002)	(4.367)	(0.101)
Number of Tranches in the Deal	0.141**	0.330	0.326	0.175**
	(0.021)	(1.000)	(1.586)	(0.024)
Expected Rating	-0.179**	-0.910	-2.062	-0.353**
	(0.045)	(1.000)	(2.851)	(0.027)
Average Market Share in Last 3 Months	4.408**	8.333**	1.311**	2.069**
	(0.008)	(1.000)	(0.161)	(0.256)
σ_{ξ_B}	0.000	0.000	0.000	0.001
	(0.901)	(1.000)	(1.000)	(0.990)
Complement/Substitute Effect(Γ)	1.234**	1.636**	4.830	1.333**
	(0.072)	(0.241)	(8.096)	(0.073)
σ_{AB}	0.014	0.088	0.999	0.089
	(0.541)	(1.000)	(1.000)	(0.165)
Log-likelihood	85170	86279	86276	74999
Estimation Bound	[-10, 10]	[-20, 20]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Starting Point	Self-picked	Random	Random	Random
Number of Obs. in Selection Stage	5914	5914	5914	5914
Number of Obs. in Rating Stage	40998	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. A coefficient with a star means it is significant at 90% confidence level, with two star means it is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process. Variables σ_{ξ_A} and σ_{ξ_B} are variance, and then always have lower bound zero.
Starting point - "Random" means the starting point in optimization process is generated by a uniform distribution between 0 and 1. "Self-specified" means the starting point are picked by myself based on their economic intuition.

Table 11: Estimate Two Stages as a Whole - with Reduced Parameters

	(1)	(2)	(3)	(4)
Rating Stage				
Peer Ratings	-0.953**	-0.953**	-0.953**	-0.937**
	(0.004)	(0.006)	(0.004)	(0.004)
Original Balance	-0.112**	-0.112**	-0.112**	-0.111**
	(0.013)	(0.023)	(0.012)	(0.009)
Original Support	-9.493**	-9.492**	-9.493**	-10.141**
	(0.997)	(0.940)	(1.159)	(1.504)
Number of Tranches in the Deal	-0.013	-0.013	-0.013	0.008
	(0.010)	(0.026)	(0.008)	(0.012)
Average Market Share in Last 3 Months	-2.806**	-2.807**	-2.806**	-2.840**
	(0.696)	(0.626)	(0.621)	(0.637)
Junior × 90-day plus Delinquency Rate				0.998
				(0.192)
Mezzanine × 90-day plus Delinquency Rate				0.023**
				(0.008)
Senior × 90-day plus Delinquency Rate				0.034**
				(0.012)
Selection Stage				
Moody's				
Original Balance	0.027**	0.027**	0.027**	0.028**
	(0.010)	(0.008)	(0.010)	(0.009)
Original Support	9.021**	9.017**	9.018**	10.466**
	(1.349)	(1.879)	(1.232)	(1.504)
Number of Tranches in the Deal	0.056**	0.056	0.056**	0.058**
	(0.006)	(0.071)	(0.008)	(0.008)
Expected Rating	-0.076**	-0.076	-0.076**	-0.077**
	(0.023)	(0.133)	(0.021)	(0.031)
Average Market Share in Last 3 Months	1.809**	1.808**	1.808**	2.448**
	(0.202)	(0.577)	(0.398)	(0.217)
σ_{ξ_A}	0.000	0.000	0.000	0.000
	(1.000)	(1.000)	(0.990)	(1.073)
S&P's				
Original Balance	0.050	0.049	0.049	0.045
	(0.048)	(0.597)	(0.052)	(0.043)
Original Support	-2.018	-2.027	-2.031	-2.402
	(1.834)	(1.652)	(2.703)	(2.541)
Number of Tranches in the Deal	0.141**	0.141	0.141**	0.175**
	(0.026)	(0.288)	(0.025)	(0.024)
Expected Rating	-0.177**	-0.178	-0.178**	-0.353**
	(0.078)	(0.642)	(0.065)	(0.082)
Average Market Share in Last 3 Months	4.395**	4.404	4.407**	2.075
	(1.215)	(5.289)	(1.536)	(1.778)
σ_{ξ_B}	0.000	0.000	0.000	0.000
	(1.000)	(1.000)	(0.990)	(0.994)
Complement/Substitute Effect(Γ)	1.236**	1.233**	1.234**	1.304**
	(0.140)	(0.588)	(1.000)	(0.224)
σ_{AB}	0.010	0.014	0.014	0.013
	(0.134)	(1.324)	(0.337)	(0.016)
Log-likelihood	85365	85296	85296	75124
Estimation Bound	[-10, 10]	[-20, 20]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Starting Point	Self-picked	Random	Random	Random
Number of Obs. in Selection Stage	5914	5914	5914	5914
Number of Obs. in Rating Stage	40998	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. A coefficient with a star means it is significant at 90% confidence level, with two star means it is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process. Variables σ_{ξ_A} and σ_{ξ_B} are variance, and then always have lower bound zero.
Starting point - "Random" means the starting point in optimization process is generated by a uniform distribution between 0 and 1. "Self-specified" means the starting point are picked by myself based on their economic intuition.

Table 12: Estimate Two Stages Separately - with Full Parameters

	(1)	(2)	(3)	(4)
Rating Equation				
Moody's				
Peer Ratings	-0.946**	-0.946**	-0.946**	-0.927**
	(0.006)	(0.005)	(0.007)	(0.005)
Original Balance	-0.143**	-0.143**	-0.143**	-0.092**
	(0.019)	(0.022)	(0.021)	(0.028)
Original Support	-13.044**	-13.043**	-13.044**	-24.488**
	(1.433)	(2.814)	(4.327)	(2.497)
Number of Tranches in the Deal	-0.052**	-0.052**	-0.052**	0.012
	(0.014)	(0.019)	(0.021)	(0.027)
Average Market Share in Last 3 Months	0.735	0.734	0.734	-2.816
	(0.893)	(0.832)	(0.851)	(1.171)
Junior × 90-day plus Delinquency Rate				0.656**
				(0.076)
Mezzanine × 90-day plus Delinquency Rate				0.013
				(0.027)
Senior × 90-day plus Delinquency Rate				-3.609**
				(1.393)
S&P				
Peer Ratings	-0.960**	-0.963**	-0.960**	-0.948**
	(0.005)	(0.005)	(0.006)	(0.006)
Original Balance	-0.052**	-0.095**	-0.051**	-0.087**
	(0.018)	(0.015)	(0.022)	(0.014)
Original Support	-2.367*	-3.979**	-2.439*	-2.374
	(1.369)	(1.332)	(1.393)	(1.701)
Number of Tranches in the Deal	0.016	0.007	0.016	0.021*
	(0.012)	(0.012)	(0.011)	(0.013)
Average Market Share in Last 3 Months	1.050	1.361	1.095	0.996
	(1.295)	(1.532)	(1.015)	(0.963)
Junior × 90-day plus Delinquency Rate				0.019
				(0.211)
Mezzanine × 90-day plus Delinquency Rate				0.029**
				(0.008)
Senior × 90-day plus Delinquency Rate				0.0516*
				(0.029)
Selection Equation				
Moody's				
Original Balance	0.026**	0.026**	0.027**	0.028**
	(0.008)	(0.009)	(0.010)	(0.009)
Original Support	9.155**	9.155**	8.938**	9.807**
	(1.022)	(1.353)	(0.851)	(1.530)
Number of Tranches in the Deal	0.049**	0.049**	0.059**	0.060**
	(0.006)	(0.006)	(0.007)	(0.011)
Expected Rating	-0.064**	-0.064**	-0.083**	-0.087**
	(0.017)	(0.025)	(0.014)	(0.029)
Average Market Share in Last 3 Months	1.884**	1.884**	1.769**	2.432**
	(0.262)	(0.351)	(0.326)	(0.255)
σ_{ξ_A}	0.000	0.000	0.000	0.000
	(0.658)	(0.922)	(0.974)	(0.945)
S&P's				
Original Balance	0.553	1.021	0.046	0.040
	(0.993)	(0.997)	(0.047)	(0.900)
Original Support	0.672	0.318	-2.012	-4.430**
	(1.000)	(1.000)	(2.623)	(1.696)
Number of Tranches in the Deal	1.163	9.324**	0.141**	0.170**
	(0.998)	(0.403)	(0.024)	(0.029)
Expected Rating	1.303	3.709	-0.196**	-0.354**
	(0.997)	(0.923)	(0.062)	(0.071)
Average Market Share in Last 3 Months	1.006	2.092**	4.367**	2.212
	(0.992)	(0.955)	(1.841)	(1.913)
σ_{ξ_B}	0.315	1.500	0.000	0.769
	(1.000)	(0.923)	(0.661)	(0.610)
Complement/Substitute Effect(Γ)	1.113**	1.469**	1.530**	1.099**
	(0.382)	(0.701)	(0.105)	(0.127)
σ_{AB}	0.787	0.500	0.315	0.336**
	(0.998)	(0.945)	(0.999)	(0.145)
Log-likelihood in first stage:	2525	2524	1412	1213
Log-likelihood in second stage:	85850	83763	85848	75475
Log-likelihood of the whole model:	88375	86287	87260	76688
Estimation Bound in first stage:	[-20, 20]	[-∞ +∞]	[-∞ +∞]	[-∞ +∞]
Estimation Bound in second stage:	[-∞ +∞]	[-∞ +∞]	[-∞ +∞]	[-∞ +∞]
Starting Points	Self-picked	Self-picked	Random	Random
Number of Obs. in Selection Stage	5914	5914	5914	5914
Number of Obs. in Rating Stage	40998	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. A coefficient with a star means it is significant at 90% confidence level, with two star means it is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process. Variables σ_{ξ_A} and σ_{ξ_B} are variance, and then always have lower bound zero.
Starting point - "Random" means the starting point in optimization process is generated by a uniform distribution between 0 and 1. "Not Random" means the starting point are picked by myself based on their economic intuition.

Table 13: Estimate Two Stages Separately - with Reduced Parameters

	(1)	(2)	(3)	(4)
Rating Equation				
Peer Ratings	-0.953**	-0.953**	-0.953**	-0.937**
	(0.004)	(0.004)	(0.004)	(0.004)
Original Balance	-0.112**	-0.113**	-0.112**	-0.111**
	(0.014)	(0.012)	(0.012)	(0.014)
Original Support	-9.494**	-9.465**	-9.493**	-10.105**
	(1.245)	(1.003)	(0.363)	(1.511)
Number of Tranches in the Deal	-0.013	-0.013	-0.013	0.009
	(0.010)	(0.009)	(0.008)	(0.011)
Average Market Share in Last 3 Months	-2.810**	-2.809**	-2.810**	-2.854**
	(0.711)	(0.608)	(0.593)	(0.668)
Junior $\times$ 90-day plus Delinquency Rate				4.828**
				(1.522)
Mezzanine $\times$ 90-day plus Delinquency Rate				0.023**
				(0.011)
Senior $\times$ 90-day plus Delinquency Rate				0.035**
				(0.013)
Selection Equation				
Moody's				
Original Balance	0.027**	0.027**	0.027**	0.028**
	(0.007)	(0.007)	(0.031)	(0.009)
Original Support	8.950**	9.155**	8.944**	10.771**
	(1.021)	(1.022)	(2.652)	(1.634)
Number of Tranches in the Deal	0.059	0.049**	0.059	0.047**
	(0.040)	(0.006)	(0.043)	(0.008)
Expected Rating	-0.082**	-0.064**	-0.082	-0.065**
	(0.012)	(0.017)	(0.297)	(0.029)
Average Market Share in Last 3 Months	1.775**	1.884**	1.775**	1.841**
	(0.859)	(0.262)	(0.917)	(0.265)
σ_{ξ_A}	0.000	0.000	0.000	0.000
	(0.658)	(0.433)	(1.000)	(0.996)
S&P's				
Original Balance	0.046	0.053	0.046	0.057
	(0.672)	(0.993)	(0.528)	(0.956)
Original Support	-2.009**	0.672	-2.012**	0.604
	(0.912)	(0.998)	(0.934)	(0.932)
Number of Tranches in the Deal	0.141	1.163	0.141	1.076
	(0.838)	(0.997)	(1.438)	(1.235)
Expected Rating	-0.194	1.303	-0.194	1.018
	(1.805)	(0.997)	(1.876)	(0.986)
Average Market Share in Last 3 Months	4.378**	1.006	4.374**	0.746
	(2.923)	(0.998)	(2.786)	(0.596)
σ_{ξ_B}	0.000	0.315	0.000	0.325
	(0.999)	(0.999)	(0.733)	(0.999)
Complement/Substitute Effect(Γ)	1.500**	1.112**	1.500**	0.996**
	(0.680)	(0.383)	(0.692)	(0.503)
σ_{AB}	0.000	0.787	0.000	0.816
	(0.999)	(0.999)	(3.039)	(0.745)
Log-likelihood in first stage	1413	2525	1413	2003
Log-likelihood in second stage	83882	83883	83883	73914
Log-likelihood of the whole model:	85295	86408	85296	75917
Estimation Bound in first stage:	[-10 10]	[-20 20]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Estimation Bound in second stage:	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Starting Points	Self-picked	Random	Random	Random
Number of Obs. in Selection Stage	5914	5914	5914	5914
Number of Obs. in Rating Stage	40998	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. A coefficient with a star means it is significant at 90% confidence level, with two star means it is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process. Variables σ_{ξ_A} and σ_{ξ_B} are variance, and then always have lower bound zero.
Starting point - "Random" means the starting point in optimization process is generated by a uniform distribution between 0 and 1. "Not Random" means the starting point are picked by myself based on their economic intuition.

Table 14: Estimate Two Stages as a Whole - with Full Parameters - Modified Model

	(1)	(2)	(3)
Rating Equation			
Moody's			
Peer Ratings	-0.946**	-0.946**	-0.929**
	(0.007)	(0.006)	(0.006)
Original Balance	-0.148**	-0.142**	-0.144**
	(0.023)	(0.019)	(0.025)
Original Support	-10.523**	-12.945**	-17.298**
	(1.143)	(1.821)	(2.104)
Number of Tranches in the Deal	-0.050**	-0.052**	-0.015
	(0.018)	(0.017)	(0.017)
Average Market Share in Last 3 Months	1.882**	1.848**	2.527**
	(0.336)	(0.389)	(0.140)
Junior $\times$ 90-day plus Delinquency Rate			2.912*
			(1.675)
Mezzanine $\times$ 90-day plus Delinquency Rate			0.007
			(0.018)
Senior $\times$ 90-day plus Delinquency Rate			0.016
			(0.018)
S&P			
Peer Ratings	-0.963**	-0.963**	-0.948**
	(0.006)	(0.006)	(0.006)
Original Balance	-0.090**	-0.090**	-0.088**
	(0.017)	(0.016)	(0.016)
Original Support	-4.391**	-4.387**	-4.014*
	(1.539)	(1.563)	(1.522)
Number of Tranches in the Deal	0.007	0.007	0.019*
	(0.012)	(0.010)	(0.011)
Average Market Share in Last 3 Months	4.155	8.797**	0.370**
	(5.341)	(0.183)	(0.065)
d	0.371**	0.189	0.090
	(0.054)	(0.124)	(0.104)
Junior $\times$ 90-day plus Delinquency Rate			1.029
			(0.688)
Mezzanine $\times$ 90-day plus Delinquency Rate			0.030**
			(0.008)
Senior $\times$ 90-day plus Delinquency Rate			0.044**
			(0.018)
Selection Equation			
Moody's			
Original Balance	0.026**	0.027**	0.028**
	(0.009)	(0.009)	(0.009)
Original Support	9.157**	9.018**	10.852**
	(1.412)	(1.816)	(1.432)
Number of Tranches in the Deal	0.049**	0.056**	0.046**
	(0.008)	(0.006)	(0.008)
Expected Rating	-0.064**	-0.076**	-0.052**
	(0.036)	(0.030)	(0.026)
Average Market Share in Last 3 Months	1.882**	1.848**	2.527**
	(0.336)	(0.389)	(0.140)
σ_{ξ_A}	0.000	0.000	0.001
	(1.048)	(0.973)	(0.998)
S&P's			
Original Balance	8.079	0.049	0.079
	(10.093)	(0.060)	(0.056)
Original Support	5.908	-2.808	-4.536
	(4.880)	(3.583)	(3.691)
Number of Tranches in the Deal	6.721	0.157**	0.205**
	(11.338)	(0.031)	(0.038)
Expected Rating	-3.140	-0.193**	-0.235**
	(7.360)	(0.079)	(0.096)
Average Market Share in Last 3 Months	4.155	8.797**	0.370**
	(5.341)	(0.183)	(0.065)
σ_{ξ_B}	2.104	0.073	0.001
	(4.450)	(0.635)	(0.235)
Complement/Substitute Effect(Γ)	5.698	1.318	0.769
	(7.630)	(0.341)	(0.682)
σ_{AB}	0.748	0.000	0.001
	(6.277)	(0.180)	(0.322)
Log-likelihood	86281	85173	75297
Estimation Bound	[-20, 20]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Starting Point	Self-picked	Random	Random
Number of Obs. in Selection Stage	5914	5914	5914
Number of Obs. in Rating Stage	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. A coefficient with a star means it is significant at 90% confidence level, with two star means it is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process. Variables σ_{ξ_A} and σ_{ξ_B} are variance, and then always have lower bound zero.
Starting point - "Random" means the starting point in optimization process is generated by a uniform distribution between 0 and 1. "Self-specified" means the starting point are picked by myself based on their economic intuition.

Table 15: Parameter Estimates from Models without Selection(with Full Parameters)

	(1)	(2)	(3)	(4)
Rating Equation				
Moody's				
Peer Ratings	-0.946**	-0.946**	-0.946**	-0.929**
	(0.006)	(0.009)	(0.005)	(0.006)
Original Balance	-0.143**	-0.143**	−0.143**	-0.144**
	(0.022)	(0.021)	(0.021)	(0.014)
Original Support	-13.046**	-13.043**	-13.043**	-17.255**
	(1.567)	(1.541)	(1.715)	(1.019)
Number of Tranches in the Deal	-0.052**	-0.052**	-0.052**	-0.016
	(0.015)	(0.015)	(0.017)	(0.018)
Average Market Share in Last 3 Months	0.734	0.734	0.735	0.071
	(0.916)	(0.979)	(3.445)	(0.442)
Junior × 90-day plus Delinquency Rate				0.105
				(0.635)
Mezzanine × 90-day plus Delinquency Rate				0.006
				(0.020)
Senior × 90-day plus Delinquency Rate				0.016
				(0.007)
S&P				
Peer Ratings	-0.963**	-0.963**	-0.963**	-0.949**
	(0.006)	(0.006)	(0.009)	(0.006)
Original Balance	-0.090**	-0.051**	-0.051**	-0.088**
	(0.015)	(0.019)	(0.017)	(0.015)
Original Support	-0.389**	-2.439*	-2.438*	-3.888**
	(1.313)	(1.583)	(1.614)	(1.429)
Number of Tranches in the Deal	0.007	0.016	0.016	0.018
	(0.010)	(0.012)	(0.027)	(0.012)
Average Market Share in Last 3 Months	1.573*	1.096	1.095	1.500*
	(1.000)	(1.092)	(0.764)	(0.993)
Junior × 90-day plus Delinquency Rate				0.891**
				(0.069)
Mezzanine × 90-day plus Delinquency Rate				0.028**
				(0.008)
Senior × 90-day plus Delinquency Rate				0.045**
				(0.023)
Log-likelihood	83755	85848	85847	73785
Estimation Bound	[-30, 30]	[-30, 30]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Starting Point	Self-picked	Random	Random	Random
Number of Obs. in Rating Stage	40998	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. * indicates the coefficient is significant at 90% confidence level. ** indicates the coefficient is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process.
Starting point “Random” means the starting points in optimization process are generated from a uniform distribution between 0 and 1. “Self-picked” means the starting point are picked based on their economic intuition.

Table 16: Parameter Estimates from Models without Selection(with Reduced Parameters)

	(1)	(2)	(3)	(4)
Rating Equation				
Peer Ratings	-0.950**	-0.953**	-0.953**	-0.937**
	(0.004)	(0.004)	(0.004)	(0.004)
Original Balance	-0.069**	-0.112**	-0.112**	-0.111**
	(0.015)	(0.019)	(0.013)	(0.014)
Original Support	-10.356**	-9.482**	-9.486**	-10.150**
	(1.392)	(1.123)	(1.123)	(1.249)
Number of Tranches in the Deal	-0.001**	-0.013**	-0.013**	-0.008
	(0.015)	(0.009)	(0.009)	(0.009)
Average Market Share in Last 3 Months	-6.081**	-2.804**	-2.795**	-2.836**
	(0.993)	(0.701)	(0.655)	(0.669)
Junior $\times$ 90-day plus Delinquency Rate				7.428**
				(1.800)
Mezzanine $\times$ 90-day plus Delinquency Rate				0.023**
				(0.008)
Senior $\times$ 90-day plus Delinquency Rate				0.034**
				(0.013)
Log-likelihood	88491	83882	83882	73910
Estimation Bound	[-30, 30]	[-30, 30]	[$-\infty$ $+\infty$]	[$-\infty$ $+\infty$]
Starting Point	Self-picked	Random	Random	Random
Number of Obs. in Rating Stage	40998	40998	40998	36035

Notes: Year and Tranche Dummies are included but not reported.
Standard errors of estimates are in the parenthesis. * indicates the coefficient is significant at 90% confidence level. ** indicates the coefficient is significant at 95% confidence level.
Estimation Bound is the bound of parameters in the optimization process.
Starting point "Random" means the starting points in optimization process are generated from a uniform distribution between 0 and 1. "Self-picked" means the starting point are picked based on their economic intuition.

Table 17: Marginal Effect of Peers' Rating - Based on results of estimating two-stages as a whole)

Moody's Original Rating	Num. of Ratings	Moody's Prob. of downgrading had S&P's rating been 1 notch lower	Moody's Prob. of upgrading had S&P's rating been 1 notch higher	S&P's Original Rating	Num. of Ratings	S&P's Prob. of downgrading had Moody's rating been 1 notch lower	S&P's Prob. of upgrading had Moody's rating been 1 notch higher
1 (25^{th} percentile)	12740	1.28%		1 (25^{th} percentile)	12877	2.13%	
3 (50^{th} percentile)	2900	66.45%	10.83%	3 (50^{th} percentile)	3534	86.76%	34.45%
7 (75^{th} percentile)	2562	44.13%	7.61%	8 (75^{th} percentile)	2362	80.51%	28.16%
9	2606	45.62%	7.55%	9	2106	77.18%	20.67%
10 (investment grade)	2534	45.20%	7.24%	10 (investment grade)	1908	75.56%	21.43%
11	975	44.97%	7.43%	11	705	75.46%	17.42%
12	808	34.01%	6.01%	12	297	64.58%	10.42%
16	23	4.75%	2.69%	15	74	3.07%	3.77%
17	26	2.49%	3.02%	18	68	12.67%	3.37%

Notes:

- The simulation results in column (3), (4), (7) and (8) are based on parameters estimates in column (4) of Table 12.
- Moody's original rating 16 and 17 and S&P's original rating 15 and 18 are chosen to report since those ratings have more observations than other low ratings.

Table 18: Marginal Effect of Peers' Rating - Based on results of estimating two-stages separately)

Moody's Original Rating	Num. of Ratings	Moody's Prob. of downgrading had S&P's rating been 1 notch lower	Moody's Prob. of upgrading had S&P's rating been 1 notch higher	S&P's Original Rating	Num. of Ratings	S&P's Prob. of downgrading had Moody's rating been 1 notch lower	S&P's Prob. of upgrading had Moody's rating been 1 notch higher
1 (25^{th} percentile)	12740	0.66%		1 (25^{th} percentile)	12877	2.10%	
3 (50^{th} percentile)	2900	66.80%	10.89%	3 (50^{th} percentile)	3534	86.75%	34.49%
7 (75^{th} percentile)	2562	44.07%	7.81%	8 (75^{th} percentile)	2362	80.06%	32.61%
9	2606	45.09%	7.58%	9	2106	76.71%	20.67%
10 (investment grade)	2534	45.18%	7.27%	10 (investment grade)	1908	75.11%	21.43%
11	975	44.96%	7.46%	11	705	74.97%	16.99%
12	808	34.09%	6.20%	12	297	64.16%	10.18%
16	23	4.84%	0.76%	15	74	3.22%	0.68%
17	26	2.55%	0.40%	18	68	12.73%	3.79%

Notes:

- The simulation results in column (3), (4), (7) and (8) are based on parameters estimates in column (4) of Table 12.
- Moody's original rating 16 and 17 and S&P's original rating 15 and 18 are chosen to report since those ratings have more observations than other low ratings.