Identification and Estimation of Soft Adjustment in Structural Bond Rating Models: Before and After the Dodd-Frank Act

Zhutong Gu*[1], Yixiao Jiang[2], and Shuyang Yang[3]

[1] *Peking University, HSBC Business School*

[2] *Rutgers University, Department of Economics*

[3] *Amazon Inc.*

Abstract

Credit Rating Agencies (CRAs) adjust preliminary bond ratings with knowledge beyond publicly available information. These unobserved "soft adjustments" may reflect material non-public information and rating biases due to conflicts of interest, making certain bond characteristics endogenous. We model soft adjustments as bond-specific thresholds in a semiparametric ordered-response model and exploit ownership structures of bond-issuers to control for endogeneity. Relying on the *shift restrictions*, we develop a location estimator that is widely applicable in models of ordered dependent variables with correlated heterogenous thresholds. Using Moody's initial ratings from 2000 to 2016, we find a significant reduction of soft adjustment after the Dodd-Frank reform.

*Corresponding author, email: zgu@phbs.pku.edu.cn; address: Peking University HSBC Business School 613, Xili University Town, Shenzhen, China; Tel: +86-18617094914. The authors are especially grateful for insightful comments and advisory from Roger W. Klein. The authors also acknowledge and thank participants at various conferences and seminars for helpful discussions and suggestions. All errors are purely the authors'.

Keywords: Bond Rating, Ordered Response Model, Semiparametric Estimation, Soft Adjustment, Dodd-Frank Act

JEL Classification: C51, G24, C14

1 Introduction

Credit Rating Agencies (CRAs) have served the financial market by providing summarized information on the default risk of financial instruments since 1920s. The proper functioning of CRAs reduces information asymmetry between borrowers and lenders, and is crucial to the efficiency of financial market.

In the aftermath of the recent financial crisis, the reliability of CRAs' rating methodology, however, has been scrutinized. In fact, apart from estimating default risk with financial variables, CRAs have the discretion to adjust rating outcomes subject to their own understanding of qualitative factors beyond spreadsheet. This step is termed "soft adjustment" by Kraft (2014) because adjustments are made based on hidden and subjective factors such as the manager's ability. Given their exposure to a variety of conflicts of interest, it is unclear whether CRAs utilize the soft adjustment to reflect material nonpublic information or simply distort ratings in a subtle way.[1] As a regulatory response, one chapter of the Dodd-Frank Wall Street Reform and Consumer Protection Act (henceforce "the Dodd-Frank") is dedicated to rating agency reform through enhancing information transparency and strengthening supervision. By rigorously analyzing a structural bond rating model, this paper contributes to the understanding of (i) how CRAs assign ratings based on public and private information, and (ii) the impact of the recent Dodd-Frank Act on the form of soft adjustment for publicly listed CRAs.

[1] There has been an extensive literature focusing on the conflicts of interest in credit ratings (Kedia et al., 2016; Becker and Milbourn, 2011; Jiang et al., 2012). Several channels, through which rating inflation can occur, are examined and documented by past studies. For example, public firms are operated under intensive pressure to grow and increase profits (Bogle, 2005), which motivates CRAs to report inflated rating in order to retain repeated customers for rating fees under the current issuer-pay business model (Cornaggia and Cornaggia, 2013; Jiang et al., 2012).

We examine the rating process for corporate bond ratings in an innovative econometric framework. To be specific, we model the aforementioned soft adjustment as *firm-specific* thresholds in an ordered-response model, in which a bond will be assigned to a rating category if its latent default index is between the corresponding thresholds. Extant literature on bond ratings (Kaplan and Urwitz, 1979; Blume et al., 1998; Huang et al., 2004, etc) often requires the thresholds to be fixed parameters, which leaves the differences in *ex post* ratings completely attributable to the idiosyncratic rating errors for observationally identical bonds. In contrast, firm-specific threshold echoes the aforementioned "soft" adjustment, because firms with identical fundamentals can wind up with different ratings due to adjustment on the thresholds. Of interests are both the latent index coefficients, which capture the relative importance of different risk predictors, and thresholds parameters, which convey information on the degree of soft adjustment.

We note that in this context, soft adjustment may stem from two sources: unobserved heterogeneity and conflicts of interest. Both of them may cause mechanical correlation between the firm-specific thresholds and regressors, leading to the problem of endogeneity. Our paper contributes to the literature of location estimators of ordered response models by allowing endogenous regressors and correlated thresholds (Manski, 1985; Horowitz, 1992; Lewbel, 1997, 2000; Klein and Sherman, 2002, etc). In particular, we demonstrate that under an additive separability condition, a bond issuer's "connectedness" with CRA's can be exploited as an efficient control for endogenity. Our identification strategy on index parameters is therefore reminiscent of the control function approach (Blundell and Powell, 2004; Florens et al., 2008; Imbens, 2007; Imbens and Newey, 2009; Kasy, 2011; Masten and Torgovitsky, 2014, etc). To quantify the soft adjustment, we focus on the average threshold conditional on the aforementioned measure of "connectedness". The identification strategy exploits a special property of the rating probability functions termed *conditional shift restrictions*, which is a generalization of Klein and Sherman (2002) to models with endogenous predictors. The regression coefficients are estimated by semiparametric pseudo maximum likelihood, combined with a grid-search algorithm to estimate conditional mean thresholds. Since there is no numerical optimization involved in the second stage, the proposed method is computationally attractive.

We estimate the proposed model with 11,134 initial bond ratings issued by Moody's from 2000 to 2016, with the sample split by the enactment of the Dodd-Frank Act in 2010. Moody's became a public firm in 2000, with over 300 shareholders every quarter since then. One conflict that resurfaced recently concerns the ownership structure of CRAs: publicly traded CRAs may bias ratings towards issuers that are invested by their own shareholders. This issue was first noted by Kedia et al. (2014) and examined in Kedia et al. (2016); Jiang (2017). However, none of the above papers explicitly considers the relationship between soft adjustment and rating bias. In contrast, we assess the change of soft adjustment after the passage of Dodd-Frank and examine the relationship between shareholding structures and soft adjustments, accounting for the fact that bond issuers may choose issue characteristics (such as issue amount, subordination status) endogenously based on perceived favorable treatment and private soft information.

The estimation results suggest several noticeable changes in terms of Moody's rating methodology before and after the Dodd-Frank. First, issuing amount and profitability has gained more weights in CRAs' discretion of hard information, implying that Moody's has become more stringent towards low-profit and high-debt issuers. Second, there is a significant drop in the dispersion of soft adjustment for all categories after the reform. This provides evidences in support of the effectiveness of the Dodd-Frank Act in credit rating industry, at least in terms of the effort to reduce information opacity. Third, by examining the pattern of threshold parameters, we find that it has become more difficult for a bond to be rated as investment grade bonds on average, and that Moody's has been more stringent towards issuers that are related with Moody's large shareholders. These findings suggest Moody's has become more conservative and sensitive to conflicts of interest after the Dodd-Frank.

The rest of this paper is structured as follows. In Section 2, we consider the stylized rating process for corporate bonds with soft adjustment and discusses identifying strategies of the hidden soft adjustment along with the assumptions needed. In Section 3, we propose a two-stage semiparametric index and location estimator. Section 4 provides the background of the U.S. credit rating industry and presents our data. Empirical results are in Section 5. Finally, Section 6 concludes this paper.

2 A Stylized Bond Rating Model with Soft Adjustment

Uncovering the "blackbox" of the rating methodology used by CRAs has always been a pursuit of financial regulators, academic researchers and business practitioners. Let $Y_i \in \{0, 1, \cdots, J-1\}$ be a discrete ordinal credit rating for bond i. To fix idea, we characterize the rating process as:

$$Y_i = \sum_{j=0}^{J-1} j\{T_{j-1,i} < V_{0i} \leq T_{ji}\}, \;\; j \in \{0, 1, \cdots, J-1\} \tag{2.1}$$

wherein a bond's latent "default risk index" is driven by observed firm and bond characteristics $X_i \equiv (X_i'^F, X_i'^B)'$ as a single-index, i.e. $V_{0i} \equiv X_i\beta_0$; $\boldsymbol{T_i} = (T_{0i}, \cdots, T_{J-1,i})$ is a vector of *bond-specific* thresholds that partitions the risk index into different rating categories.

Importantly, allowing the thresholds to be bond-specific echoes the idea of "soft adjustment" (Kraft, 2014).[2] That is, by perturbing $\boldsymbol{T_i}$, the CRA is able to assign different ratings to bonds that have identical financial characteristics, reflecting certain qualitative adjustments. Extant credit rating models often restrict thresholds to be constant plus a pure random error; this assumption, however, rules out soft adjustment, leaving the differences in *ex post* ratings completely attributed to the idiosyncratic rating errors for otherwise identical bonds.

[2] According to Kraft (2014) and Petersen (2004), credit ratings involve both "hard" and "soft" information. The CRA first constructs a default risk index using publicly available quantitative information from issuers' financial statements. This generates the relatively objective "hard" information. Next, the CRA conducts a subjective "soft" adjustment based on other qualitative factors and then finally release a categorical rating to the public given its internal criteria.

Of interests are the estimation of the index coefficients β_0 and recovering patterns of the bond-specific threshold $\boldsymbol{T}_i$. However, when $\boldsymbol{T}_i$ is interpreted as soft adjustments, the correlation between $\boldsymbol{T_i}$ and regressors requires new identification strategies to be developed. We believe such correlation may arise from two sources: *unobserved heterogeneity* and the *conflicts of interest*. The unobserved heterogeneity reflects qualitative information about the underlying bond that is otherwise not reflected in X_i and this information is available to both the bond-issuing firm and the CRA but not the researcher.[3] Therefore, the bond-specific threshold should correlate with X_i due to the mechanical correlation between a bond's observed and unobserved characteristics. In the online appendices, we present a simple bond issuing model highlighting the fact that some bond and firm characteristics, X_i, are likely to be endogenously selected in the presence of unobserved heterogeneity.

Other than reflecting CRA's private information, the soft adjustment may also be due to conflicts of interest, especially so in our context that bond issuers and the CRA are connected through common shareholders, and that the CRA may cater to the interest of its shareholders by assigning inflated ratings. Endogeneity arises in this context because the issuer can contemplate on the choices of characteristics by forming an expectation of the soft adjustment given their own private relationship with the CRA. For instance, a better connection with the CRA could induce riskier and more audacious issuance due to the expectation of receiving a upward rating adjustment.[4] Estimates will be generally inconsistent if the induced endogeneity is not appropriately accounted for.

Before discussing our identification strategy in detail, we make several observations. First, the soft adjustment, if identified, might serve the purpose of uncovering the "blackbox" of rating models, increasing the transparency of rating methodology and improving the

[3]For example, the unobserved firm characteristics could encompass the managerial efficiency, organizational productivity or other qualitative financial risk related information beyond spreadsheet. Moreover, this firm heterogeneity is very likely to be time-varying, so adding company fixed effects might not fully solve the problem.

[4]To illustrate, consider the following example: one bond characteristics that affects credit ratings is *issuing size*, namely how large the debt issuance is. If the bond issuer has prior information about $\boldsymbol{T_i}$, say knowing that the CRA will be more lenient, it may strategically choose to issue more debt since a better credit rating could lower the borrowing interest rate. See online appendices for the example.

predictive power of current models. Second, this econometric framework can be utilized to investigate whether the rating process is impaired by conflicts of interest. By estimating the systematic pattern of T_{ji}, one can empirically test whether the CRA has consistent rating criteria towards all bond issuers.

2.1 Identification via Shareholding Structure

It should also be noted that our model is semiparametric in the sense that no distributional assumption is made about the thresholds T_{ji}. And we consider this necessary especially given the complex structure and meanings of the soft adjustment across bonds and categories. For this reason, the index parameter vector β_0 can be identified only up to location and scale. Suppose $\beta_0 \equiv (\beta_{10}, \beta_{20}, \cdots, \beta_{d0})' \in \mathbb{R}^{d-1}$ is the coefficient vector that is conformable to the $d-$dimensional bond and firm characteristics X_i. We let $\beta_{10} = 1$ for a continuous variable X_{1i}, e.g. the log of total asset in our empirical analysis, and denote the identifiable index by $V_{0i} \equiv X_{1i} + \widetilde{X}_i'\beta_0$. Under this normalization, β_0 becomes the relative contribution of each characteristic with respect to that of X_{1i}. A sufficient condition is $\det(\widetilde{X}_i'\widetilde{X}_i) > 0$ with X_{1i} being a continuous variable and $\widetilde{X}_i \equiv (X_{2i}, \cdots, X_{di})$.

Preceding the discussion of identification, we rely on a control function approach to handle endogenous variables that are correlated with structural thresholds.[5] To be specific, let R_i be a vector capturing the whole connectedness between the CRA and a bond issuer through common shareholders, e.g. each common shareholder's identity, investment stake in the CRA and the bond-issuing firm i, etc.

We impose an additive separable structure on thresholds.

A-I.1 Additive Separability. $T_{ji} = \delta_j(R_i) + u_{ij}, \ \ \forall j, \quad$ *where* $u_{ij} \perp (X_i, R_i)$.

A-I.1 basically conveys that each threshold can be decomposed into two additive terms, i.e. a category-specific component, $\delta_j(\cdot)$ and a orthogonal random shock u_{ij}. The form of $\delta_j(\cdot)$

[5] The control function approach is frequently employed in nonseparable models with endogeneity (Blundell and Powell, 2004; Florens et al., 2008; Imbens, 2007; Imbens and Newey, 2009; Kasy, 2011; Maurer et al., 2011; Masten and Torgovitsky, 2014; Hoderlein and Sherman, 2015, etc).

needs not to be specified. The former component reflects the heterogeneous soft adjustment or category effect of interest, while the disturbance represents calculation errors of the CRA or pure noise. Under A-I.1, it implies that the soft adjustment is equivalent to the conditional mean of threshold, i.e. $E(T_{ji}|R_i) = \delta_j(R_i)$. It also implies the independence of X_i and the thresholds T_{ji} after conditioning on R_i.[6] To see this, note that

$$\Pr(Y_i \leq j|X_i, R_i) = \Pr(-u_i \leq \delta_j(R_i) - V_{0i}|X_i, R_i) = \Pr(Y_i \leq j|V_{0i}, R_i) \tag{2.2}$$

Our usage of shareholding structure R_i as controls is motivated by the sources of soft adjustment. In the case of the *conflict-of-interest*, having stronger connections with the CRA may lead to favorable adjustment at each category. When thresholds reveal the knowledge of *unobserved heterogeneity*, our compromising assumption is that common shareholders have as much private information as the CRA. If so, any material nonpublic information about the issuer firm is “materialized” in common shareholders’ investment decisions. For example, institutional investors are more likely to invest in issuers with “better” unobserved soft quality, inducing mechanical dependence between unobserved quality and R_i. In empirical contexts, it is alway the case that both sources of adjustment are present. However, it is unnecessary to separately identify the two sources if our objective is only to consistently estimate the index parameters and back out the *ex post* individual soft adjustment.[7] For identification purpose, we have to abstract away any other form of unobserved heterogeneity beyond those that can be relayed through the shareholding structure.

In order to encompasses the entire connectedness structure between bond issuer i and the CRA through all common shareholders, R_i can be very high-dimensional. To this end, we assume that the information contained in R_i can be sufficiently summarized by a “index” of three key aggregate variables in the empirical analysis such that

$$\Pr(Y_i \leq j|X_i, R_i) = \Pr(Y_i \leq j|X_i'\beta_0, R_i'\alpha_0) \tag{2.3}$$

[6]This is also implied by a weaker condition—conditional independence, i.e. $X_i \perp \boldsymbol{T}_i|R_i$.

[7] One shortcoming of mixing together those errors is that counterfactuals in regards to the change of shareholding structure would be confounded and unclear.

Under this simplification, the space of finite-dimensional parameters also expands, e.g. $\theta_0 \equiv (\beta_0, \alpha_0) \in (\mathcal{B} \times \mathcal{A})$. Such models have been studied in Ichimura and Lee (1991). Identification of double-index parameters requires the existence of at least one continuous variable in each index and a sufficient condition precluding the composition of same variables in both indices, which our model has already satisfied.[8] To steamline the discussion, we defer the detailed description of these three variables until the empirical section. As our identification of soft adjustment does not rely on the index structure of R_i, we therefore illustrate it in a general specification.

2.2 Soft Adjustment and Conditional Shift Restrictions

Turning to the soft adjustment $\delta_j(R_i)$, the key object of interest in this paper, we show that identification of this object can be achieved by exploiting a "special" property of the rating probabilities. This property is a generalization of the "shift restriction" proposed by Klein and Sherman (2002)[9]. Besides ours, there are other thresholds or location estimators that have been considered in the binary or ordered choice literature (Manski, 1985; Horowitz, 1992; Lewbel, 1997, 2000, 2003; Chen, 1999, 2000, etc). But most of them focus on models with only exogenous variables.

We begin by introducing notations. Recall from A-I.1 that for a representative category j, the soft adjustment can be expressed as $\delta_j(r) = E(T_{ji}|R_i = r)$, e.g., the conditional expectation of the $j's$ threshold given $R_i = r$. Likewise, we define $\Delta_{j,k}(r) \equiv E(T_{ki} - T_{ji}|R_i = r)$ as the conditional mean threshold difference between categories k and j.[10] With the conditional cumulative rating probability function defined as

$$P_j(v, r) \equiv \Pr(Y_i \leq j | V_{0i} = v, R_i = r), \quad j \in \{0, 1, \cdots, J-1\}, \tag{2.4}$$

[8] As in other semiparametric models, index parameters are identified up to location and scale. Specifically, let $R_{1i} + \widetilde{R}_i' \alpha_0$, where R_{1i} is continuous and $\widetilde{R}_i \equiv (R_{2i}, R_{3i})$, with $|\det(\widetilde{R}_i \widetilde{R}_i')| > 0$.

[9] In Klein and Sherman (2002), they use those shift restrictions to estimate the relative scaled thresholds in semiparametric ordered response models. In this paper, we generalize the shift restriction technique to allow endogenous regressors with correlated thresholds.

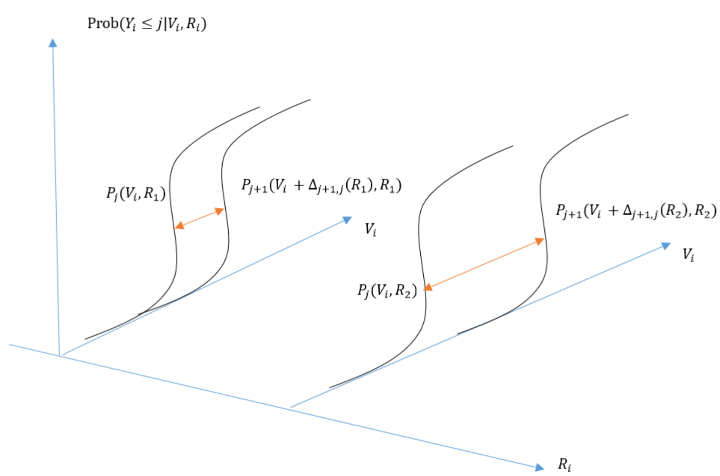
[10] A related parameter of interest, resembling the "treatment effect on the treated", is defined as $\delta_j^c(x) \equiv E(T_j|X_i = x)$ for each x. Analogously, let $\Delta_{k,j}^c(x) \equiv E(T_{ji} - T_{ki}|X_i = x)$.

the above function measures the probability of bond i being rated into category j or above given its true risk index and shareholding relationship with the CRA. Identification of soft adjustment relies critically on the implication of Proposition 1 below.[11]

Proposition 1 Conditional Shift Restriction**.** *Under Assumption A-I.1, for each* $(v, r) \in \mathbb{R} \times \mathcal{R}$*, then* $P_j(v, r) = P_k(v + \Delta_{j,k}(r), r)$*, for each* $j, k \in \{0, 1, \cdots, J-1\}$*.*

Proposition 1 introduces the conditional shift restrictions, which is a natural generalization of Klein and Sherman (2002)'s. In particular, it reveals the hidden restrictions across categories. Figure 1 depicts an example of shifts between jth-conditional probability functions and the $(j+1)$th for single-dimensional R_i. Given the index $V_{0i} = v$ and the conditioning variable $R_i = r$, one can equate $P_{j+1}(v, r)$ to $P_j(v, r)$ by increasing the index v by $\Delta_{j,k}(r)$, which is precisely the conditional mean thresholds differences. Intuitively, identification of $\Delta_{j,k}(r)$ can be achieved by equating the two probability functions $P_{j+1}(v, r)$ to $P_j(v, r)$ for a given r.

Figure 1: Conditional Shift Restrictions from $P_j(V_i, R_i)$ and $P_{j+1}(V_i + \Delta, R_i)$



[11]Note that A-I.1 is sufficient but not necessary for conditional shift restrictions. A weaker set of conditions require that the conditional distributions of u_{ij} be the same for each level j. However, we would lose the interpretation of the $\delta_{j,k}(r)$ being the soft adjustment.

More formally, $\Delta_{j,k}(r)$ can be identified by inverting $P_k(\cdot, r)$

$$\Delta_{j,k}(r) = P_k^{-1}(P_j(v,r), r) - v \tag{2.5}$$

for each r and $j, k \in \{0, 1, \cdots, J-1\}$. Formal proof of Proposition 1 is given in the appendix. Under proper support conditions, the unconditional mean of thresholds can be obtained by taking expectation, i.e. $\Delta_{j,k} = E[\Delta_{j,k}(R_i)]$.[12]

Recall that $\Delta_{j,k}(r)$ only measures the average distance between thresholds as a function of the shareholding structure R_i. In practice, of interests are often the level of thresholds themselves as they reflect the CRA's rating criteria.[13] To make comparison of thresholds possible, we assume that

A-I.2 Base level. *There exists at least one j such that $\partial\delta_j(r)/\partial r = 0$ for any $r \in \mathcal{R}$.*

In the empirical estimation, we choose $j = 0$ and assume that there is essentially no heterogeneous soft adjustment for Aaa-rated bonds, i.e. $\delta_0(r) = 0$, for any $r \in \mathcal{R}$. As a result, the soft adjustment at category $j \in \{1, \cdots, J-2\}$ can be backed out as $\Delta_{0,j}(r) = E(T_{ji} - T_{0i}|R_i = r) = \delta_k(r)$. We provide further support for our choice of the normalized category in the empirical section.

3 A Two-stage Semiparametric Estimator

In this section, we provide a two-stage semiparametric estimators for $(\theta_0, \mathbf{\Delta}(r))$ for each $r \in \mathcal{R}$, where $\mathbf{\Delta}(r) \equiv (\Delta_{0,1}(r), \Delta_{0,2}(r), \cdots, \Delta_{0,J-2}(r))'$ denotes the identified vector of threshold differences. In the first stage of estimation, we target at the index parameters, $\theta_0 \equiv (\beta_0, \alpha_0)$, up to location and scale by pseudo-maximum likelihood (ML) estimation. After obtaining estimates for $\widehat{\theta}$, one can thus construct the risk index estimator $\widehat{V}_i = X_i'\widehat{\beta}$ as well as the

[12]In case the large support condition fails, one may compute a set-average expectation: $\Delta_{j,k}(\mathcal{R}^0) = E[\Delta_{j,k}(R)|R \in \mathcal{R}^0]$, where $\mathcal{R}^0 \subset \mathcal{R}$ is a compact set of interest.

[13]For example, a bond can be rated at least category j when its latent default risk is less than T_{ji}. When the threshold point increases, $(j+1)$-rated bonds that were slightly above the threshold can now be rated as category j. This implies a loosened rating criteria.

relationship control index $R_i'\widehat{\alpha}$. In the second stage, we estimate the conditional mean thresholds (or soft adjustment), $\widehat{\Delta}(r)$ at each point $r \in \mathcal{R}$ by a grid search estimator. The grid search algorithm is attractive for its fast computing speed, as opposed to GMM and other extremum estimators.[14]The asymptotic properties of index estimators follows directly from Ichimura and Lee (1991).

3.1 First Stage: Index Estimators

Conditional Probability Function We begin by introducing the estimator of conditional rating probability function in Eq. (2.4), since they serve as the basic building blocks that will be fed into the likelihood function. Semiparametric single (and multiple) index estimators have been extensively studied in the literature. For exogenous covariates, see Manski (1985); Powell et al. (1989); Klein and Spady (1993); Ahn et al. (1996); Klein and Shen (2010), etc. For models with endogeneity, see Blundell and Powell (2004); Hoderlein and Sherman (2015), etc.

For any $(\beta, \alpha) \in \Theta$, define $V_i(\beta) \equiv X_i'\beta$ and $R_i(\alpha) \equiv R_i'\alpha$ and we suppress θ for notational simplicity whenever it is self-evident. We use the local constant kernel estimator to obtain the semiparametric conditional probabilities. In particular, the leave-one-out semiparametric estimator of the conditional probability function for $Y_i \leq j$ is used in Eq. (3.1),

$$\widehat{P}_j(i;\theta) \equiv \widehat{P}_j(V_i(\beta), R_i(\alpha)) = \frac{\sum_{l \neq i}^{N} K_h(V_l(\beta) - V_i(\beta))K_h(R_l(\alpha) - R_i(\alpha))\{Y_l \leq j\}}{\sum_{l \neq i}^{N} K_h(V_l(\beta) - V_i(\beta))K_h(R_l(\alpha) - R_i(\alpha))} \tag{3.1}$$

To seek a $\sqrt{N}$-consistent parameter estimator, we could resort to the bias reduction techniques available in the semiparametric literature to make sure that the asymptotic bias vanishes faster in the limit. In principle, one may use higher-order kernels, local smoothing and the recursive methods in Shen and Klein (2017). In practice, we only use Silverman's rule-of-thumb bandwidths as we do not find large and significant change to our results with

[14]The two-stage estimator can be combined in a single-step GMM estimator, which might lead to more efficient estimation. However, doing it in two stages would be much faster in practice when the dimension of parameter space gets large.

bias-correction techniques.[15]

Pseudo-ML Estimator Note that the double-index parameters can be solved in a semiparametric pseudo-MLE framework similar to Klein and Vella (2009); Maurer et al. (2011). It requires only one-step of optimization and can be computationally fast. Take the single index model as an example, define $\widehat{P}_{-1i}(\theta) = 0$ and $\widehat{P}_{Ji}(\theta) = 1$.

$$\widehat{\theta} = \arg\max_{\theta \in \Theta} N^{-1} \sum_{i=1}^{N} \sum_{j=0}^{J} \widehat{t}_i \{Y_i = j\} \ln \left(\widehat{P}_j(i;\theta) - \widehat{P}_{j-1}(i;\theta) \right)$$

where the trimming function estimator $\widehat{t}_i = \Pi_{k=1}^{d_X + d_R} \{\widehat{q}_{Z_k}(\tau_l) < Z_{ki} < \widehat{q}_{Z_k}(\tau_u)\}$ is the product of the indicator functions for each continuous Z_k, with fixed lower and upper quantiles τ_l and τ_u, where $Z_i = (X_i', R_i')'$. $\widehat{q}_{Z_{ik}}(\tau)$ is estimated by the empirical quantile function, $\inf\{z_k : N^{-1} \sum_{i=1}^{N} \{Z_{ki} \leq z_k\} \geq \tau\}$.

3.2 Second Stage: Conditional Mean Thresholds $\Delta(\cdot)$

The shift restrictions naturally implies an extremum-type estimator by minimizing the distance between $P_j(V_i, r)$ and $P_k(V_i + \Delta_{k,j}(r), r)$ for each $r \in \mathcal{R}$, $j \neq k$ and $j, k \in \{0, 1, \cdots, J-1\}$. For a J-supported Y_i, there are totally $\binom{J-2}{2}$ possible restrictions to choose from. In order to have a parsimonious model, we only consider the shift conditions of adjacent levels. Additional restrictions could be used to increase the efficiency and perform overidentification test.

However, in terms of computing time, the optimization needs to be done repeatedly for each value of R_i in the sample or of particular interest. This can take quite long time once the support of Y_i is large. To this end, we choose to estimate it by directly inverting the conditional probability functions following the identification condition in Eq. (2.5). Since the equality holds for each value of v, the final estimator takes the form of averaging over

[15] For the optimal bandwidth, $h = 1.06 \times std \times N^{-1/6}$.

all empirical points of $V_i(\beta)$, for $i \in \{1, 2, \cdots, N\}$.

$$\widehat{\Delta}_{j,j-1}(r) = \frac{1}{N}\sum_{i=1}^{N}\left[\widehat{P}_{j-1}^{-1}\left(\widehat{P}_j(V_i(\widehat{\beta}), r), r\right) - V_i(\widehat{\beta})\right], \quad j \in \{1, \cdots, J-1\} \tag{3.2}$$

We also choose to estimate the adjacent levels as the range of overlapping points is the largest. Without redundant information, we are left with $J-2$ restrictions and for each $r \in \mathcal{R}$. To implement it, we use a grid search algorithm which is a simple extension of the grid search estimator in Klein and Sherman (2002).[16]

One can repeat the following four steps for each $r \in \mathbb{R}$ and $j \in \{1, \cdots, J-1\}$,

1. Estimate $\widehat{P}_j(\widehat{V}_i(\widehat{\beta}), r)$ nonparametrically for each i.
2. Estimate $\widehat{P}_{j-1}(v, r)$ nonparametrically at each v over a set of grid points, $\mathcal{V}_N^g$.
3. Find the closest v such that $V_i^* = \arg\min_{v \in \mathcal{V}_N^g} |\widehat{P}_j(\widehat{V}_i(\widehat{\beta}), r) - \widehat{P}_{j-1}(v, r)|$ for each i.
4. Compute $\widehat{\Delta}_{j,j-1}(r)$ as $N^{-1}\sum_{i=1}^{N} V_i^* - V_i(\widehat{\beta})$.

For the choices of grid sets $\mathcal{V}_N^g$, it is advised that the adjacent interval distance should be smaller than $O(1/\sqrt{N})$ in order to be negligible in the limit. Provided the fast speed of the grid search algorithm, one can pick even finer grid, though empirical differences are not quite significant. The relative conditional mean threshold of level j with respect to the base level (namely $Y_i = 0$) is readily available by multiplying a lower triangular matrix A with entry equal to 1 below and along the diagonal. Let $\widehat{\Delta}^0(r) = A\widehat{\Delta}(r)$, so $\widehat{\Delta}^0(r) = (\widehat{\Delta}_{1,0}(r), \widehat{\Delta}_{2,0}(r), \cdots, \widehat{\Delta}_{J-2,0}(r))'$.

If it is innocuous to normalize the base level $\delta_0(r) = 0$ as suggested in Assumption A-I.2, then we can back out the unobserved soft adjustment starting from $j = 2$ to $j = J-1$. To do so, we first calculate the empirical control index $R_i(\widehat{\alpha}) = R_i'\widehat{\alpha}$ and then compute the relative thresholds evaluated at each $R_i(\widehat{\alpha})$. By definition, estimates of individual-bond

[16]The joint estimation of a vector of conditional thresholds using extremum-type estimators could be more efficient than our grid search methods. However, not only the computational time would substantially increase but the estimates can be sensitive to starting values.

soft adjustment at each category j would be $\widehat{\delta}_{ij} = \widehat{\Delta}_{j,0}[R_i(\widehat{\alpha})]$. In the next section, we will examine the empirical distributions of soft adjustment before and after the Dodd-Frank Act.

Note that the consistency and asymptotic normality of the finite-dimensional index parameter estimators are very standard in the semiparametric literature. As for the inference, we apply the asymptotic normality condition in Klein and Sherman (2002) to double-index models. For the soft adjustment estimator in Eq. (3.2), consistency of $\widehat{\delta}_j(r)$ would straightforwardly follow once the consistency of $\widehat{P}_j(\cdot, r)$ and $\widehat{V}_i$ are established.[17] For the inference of the thresholds estimators, we bootstrap the variances starting only from the second stage. To implement it, we repeatedly draw samples with the same number of observations over all possible $(Y_i, \widehat{V}_i, R_i)$ with replacement. Fortunately, our computational time is still relatively fast, even for some large number of bootstrapped samples.

4 Data and Context

4.1 Institutional and Regulatory Environment

As the information intermediaries of the financial system, credit rating agency's primary function is to evaluate a particular debt instrument's credit worthiness. As noticed by Cantor et al. (1994) and White (2002), the credit rating industry in the U.S. is highly concentrated: with the"Big Three" credit rating agencies controlling more than 95% of the ratings business. Moody's and Standard & Poor's (S&P) together control 80% of the global market, and Fitch Ratings controls a further 15% (Alessi et al., 2013). Of the two biggest agencies Moody's became a public firm in 2001, while Standard & Poor's is a private division of the McGraw-Hill.

Given the massive defaults of highly-rated securities during the last financial crisis, various reforms have been proposed to regulate the behaviors of CRAs. In the famous Dodd-Frank Wall Street Reform and Consumer Protection Act (Pub.L. 111203[18], H.R. 4173[19]),

[17] The point-wise consistency result is given in the online appendices.

[18] https://www.gpo.gov/fdsys/pkg/PLAW-111publ203/html/PLAW-111publ203.htm

[19] https://www.congress.gov/bill/111th-congress/house-bill/4173

an entire section is devised to improve the transparency of credit rating agencies, by means of enforcing public disclosure of credit rating methodologies, data, and etc. In subtitle C of Title IX of the amendments, it emphasizes on the “improvements to the regulation of credit rating agencies, critical *gatekeeper* in the debt market central to capital formation, investor confidence, and the efficient performance of the United States economy.” Subtitle C also cites findings of conflicts of interest and inaccuracies during the recent financial crisis contributed significantly to the mismanagement of risks by financial institutions and investors, which in turn adversely impacted the health of the United States economy.[20] The Franken-Wicker amendment to the Dodd-Frank financial reform law,[21] taking a somehow more extreme approach, suggests to use a governmental entity to assign securities to qualified ratings agencies based on capacity and expertise. Recall from the earlier discussion that rating agencies make soft adjustment based on private information. By estimating how soft adjustments evolve over time, we aim to assess the effectiveness of Dodd-Frank in enhancing information transparency.

4.2 Data and Summary Statistics

Our data derive from multiple sources. The data on the history of credit rating by Moody’s is obtained from the Mergent’s Fixed Income Securities Database (FISD). Our sampling period spans from 2001, when Moody’s went IPO, to 2016, with the enactment of Dodd-Frank in July 2010. Given the aforementioned regulatory changes, we divide the sample into two time periods by the enactment of the Dodd-Frank on July 21, 2010. To allow for possible implementation lags, we alternatively define the post Dodd-Frank period starting from the beginning of 2011. By comparing estimates of soft adjustment in both periods, we

[20]This law required the SEC to establish clear guidelines for determining which credit rating agencies qualify as Nationally Recognized Statistical Rating Organizations (NRSROs) who are required to establish, maintain, enforce and document an effective internal control structure governing the implementation of and adherence to policies, procedures, and methodologies for determining credit ratings. It also gave the SEC the power to regulate NRSRO internal processes regarding record-keeping and how they guard against conflicts of interest. See Partnoy (2009) and White (2002) for the importance of such oversight. https://www.sec.gov/spotlight/dodd-frank/creditratingagencies.shtml

[21]https://www.sec.gov/comments/4-629/4629-28.pdf

aim to examine the policy effect of Dodd-Frank Act on the credit rating outcomes. In the pre Dodd-Frank period, a crisis dummy is created to capture the financial crisis effect from 2007 to 2010.

We exclude government bonds and retain all initial ratings on bonds issued by firms covered in both Center for Research in Security Prices (CRSP) and Compustat, leaving us with a final sample of 11,134 initial bonds issued by 1462 firms. There are 2,540 observations in the crisis period, accounting for 38.4% of the total before the Dodd-Frank. The distribution of ratings are presented in Table 1 by years, with Aaa being the highest credit category and C the lowest. Figure 2 compares the rating distribution before and after the Dodd-Frank. Noticeably, Moody's has become more conservative within the investment grade class (Baa and above), as the share of A and Aa bonds have declined while the share of Baa bonds rose abruptly.

It worths to pointing out that we choose to focus exclusively on Moody's rating process because of the agency's unique exposure to conflicts of interest. Among the "big three" rating agencies on the market, Fitch is a private firm; Being a private division under McGraw-Hill, the conflicts-of-interest effect is complicated and indirect for S &P. Focusing exclusively on Moody's ratings, however, might lead to sample selection bias. However, we do not view this as a serious problem in our sample because most of bonds issued by public firms have been rated by at least two firms and their ratings are mostly matched after being converted to the same standard (70% of the ratings assigned by S&P and Moody's differ by at most one notch). This implies that agency heterogeneity does not play quite a fundamental role in the rating methodology to some extent.

4.2.1 Firm and Bond Characteristics

Using data from quarterly Compustat-CRSP merged database and FISD, we construct a sequence of predictors for credit ratings mentioned in the bond rating literature (Pinches and Mingo, 1973; Kaplan and Urwitz, 1979; Blume et al., 1998; Jiang et al., 2012; Campbell and Taksler, 2003, etc). To construct these variables, short-term and long-term debt for each

Table 1: Moody's Rating Outcomes by Year

Year	Aaa	Aa	A	Baa	Ba	B	C	Total
2001	10	45	171	221	122	96	11	676
2002	1	78	146	223	85	107	7	647
2003	9	112	155	219	131	174	32	832
2004	3	85	95	177	103	160	18	641
2005	6	118	115	161	94	92	15	601
2006	4	164	163	195	63	68	24	681
2007	9	238	332	167	55	75	13	889
2008	2	110	156	143	29	12	4	456
2009	3	35	129	230	89	104	13	603
2010	7	53	105	183	93	125	26	592
2011	10	38	142	226	41	98	17	572
2012	3	43	166	289	93	134	25	753
2013	12	58	181	325	109	117	36	838
2014	8	37	144	324	94	102	25	734
2015	20	35	218	399	90	68	10	840
2016	26	59	192	347	81	68	6	779
Total	133	1,308	2,610	3,829	1,372	1,600	282	11,134

Note: Subtiers are aggregated together into general tiers, e.g. A consists those rated as A1, A2 and A3.

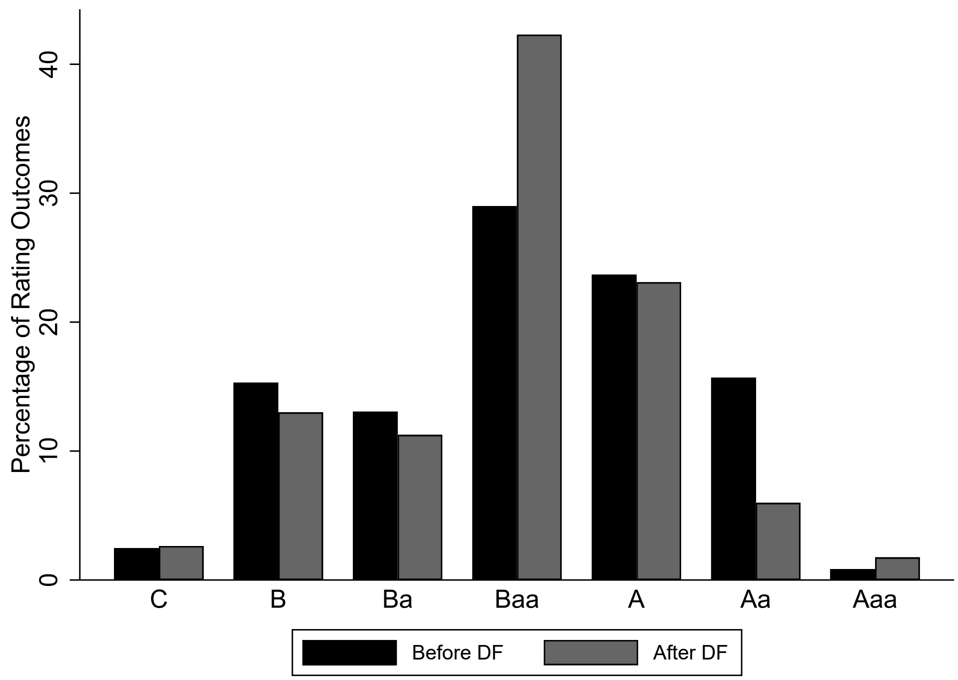
bond issuers are from quarterly Compustat-CRSP merged dataset. [22] The end of quarter stock price data and number of shares outstanding data are also taken from Compustat-CRSP. All financial ratios are computed using a 5-year arithmetic average of the annual ratios, as Kaplan and Urwitz (1979) points out that bond raters might look beyond a single year's data to avoid temporary anomalies.

The selected predictors consists of firm characteristics (1)-(4) and bond characteristics (5)-(6) as follows: (1) ASSET: denotes issuer size, defined as the value of the firm's total asset. (2) LEVERAGE: denotes firm leverage, defined as the ratio of long-term debt to total assets. (3) PROFIT: denotes operating performance, defined as operating income before depreciation divided by sales. (4) CVTA: denotes asset stability, defined as the variance of the firm's total asset in the year prior.[23] (5) OFFAMT: denotes the offering amount, defined as the par value of the bond issued. (6) SENIOR: denotes subordination status, which a

[22]Short-term debt is estimated as the larger of Compustat items 118 ("Debt in current libabilities") and 224 ("Total current liability"). Long-term debt is taken from item 119 ("Total long-term liability").

[23]The definition here follows that from (Kedia et al., 2016).

Figure 2: Rating Outcome Distributions Before and After the Dodd-Frank Act

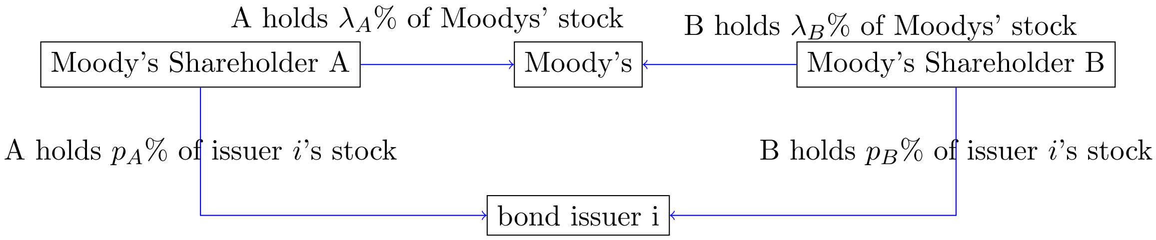


Note: 1. DF is the Dodd-Frank indicator that equals 1 for those after 2010 and 0 if before 2010.

dummy variable equals to one if the bond is a senior bond and 0 otherwise. We take log on both sizing variables (OFFAMT, ASSET) to make all covariates roughly have the same scale as their differences in denominations can be potentially large. A summary statistics of the ratings and explanatory variables can be found in the upper panel of Table 2. As motivated in the the behavioral framework, LEVERAGE, OFFAMT and SENIOR are likely to be endogenous as firms might issue more debt when they "foresee" a chance of higher ratings.

4.2.2 The Shareholding Relations

Recall from Section 2 that we need a vector R_i to characterize the shareholding relationship between Moody's and each bond issuer i, so we can address the endogeneity problem by conditioning on this control vector R_i. We use three variables to jointly capture the shareholding relationship. That is, $R_i \equiv \{\text{Mshare}_i, \text{Fshare}_i, \text{LargeSH}_i\}$. To convey some intuition on the definition of these variables and why they are selected, consider a bond issuer i that is jointly invested by two shareholders of Moodys, A and B, as described below:[24]



The shareholding relationship between Moody's and bond issuer i, in a sense, can be characterized by both *the importance of shareholders to Moody's* (captured by the $\lambda's$) and *the importance of bond issuer i to the shareholders* (captured by the $p's$)[25]. To be more precise, we aggregate the two shareholding percentage measure λ and p across all common shareholders to approximate bond issuer i's overall ownership interaction with Moody's (in this illustrative example, namely $\lambda_A\% + \lambda_B\%$ and $p_A\% + p_B\%$, respectively). Extending to the case with M_i common shareholders, we define:

$$Mshare_i \equiv \sum_{j=1}^{M_i} \lambda_j\%, \quad Fshare_i \equiv \sum_{j=1}^{M_i} p_j\%,$$

to capture the importance of shareholders to Moody's and the importance of bond issuer i to the shareholders. In addition, to highlight the individual shareholder's influence, we define:

$$\text{largeSH}_i \equiv 1\{\text{issuer } i \text{ is invested by at least one large shareholder of Moody's}\}$$

where $1\{E\}$ takes value one if E is true and zero otherwise. In particular, "large" shareholder are those who own at least 5% of Moodys' stock. The significant influence of large shareholders is also documented in Kedia et al. (2016).

The descriptive statistics for these three measures are presented in the lower panel of Table 2. In general, a larger Mshare, Fshare or largeSH indicates a stronger connection with

[24] However, Moodys could have shareholders who do not invest in the bond issuer i at all.

[25] This characterization is enlightened by Kedia et al. (2016), in which the authors find Moody's has an upward bias towards issuers that are large investees or subsidiary firms of its large shareholders.

the CRA. Mshare and largeSH provide the only channel for the conflict-of-interest to impact ratings, with the magnitude being mediated by the level of Fshare. In contrast, Fshare is supposed to pick up the unobserved bond or firm quality that is also contained in the CRA's soft adjustment. By assumptions, after conditioning Mshare, Fshare and largeSH, issuers should bear no more soft information related to issuing decisions.

As a robustness check, we have also experimented with other measures such as number of common shareholders, number of influential shareholders, number of bonds rated before as well as the weighted shares in common shareholders' portfolio. Moreover, we checked the quadratic and cubic functional forms. However, inclusion of additional variables or higher order terms increases the collinearity of the control index and lead to nonsensical coefficient estimates. Therefore, we stick to the above parsimonious specification.

Table 2: Descriptive Statistics

	Before DF: 2000-2010				After DF: 2011-2016			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
	Issuer and Bond Financial Characteristics							
ASSET	161.301	325.751	0.078	3065.556	166.851	446.162	0.175	2601.652
CVTA	0.181	0.146	0.000	1.342	0.169	0.175	0.008	1.504
LEVERAGE	0.259	0.172	0.000	1.283	0.260	0.156	0.000	1.021
PROFIT	0.028	0.058	-0.682	0.503	0.046	0.057	-0.366	0.264
OFFAMT	595.5	3860.1	0.0	250000.0	774.3	2017.2	0.0	100000.0
SENIOR	0.823	0.381	0	1	0.908	0.289	0	1
CRISIS	0.384	0.486	0	1	0	0	0	0
	Common Shareholder Information							
Mshare	44.896	15.468	0	98.794	45.808	15.073	0	89.296
Fshare	46.508	17.066	0	93.204	45.909	17.724	0	99.528
largeSH	0.620	0.485	0	1	0.898	0.302	0	1
Obs.N	6,618				4,516			

Note: 1. ASSET and OFFERAMT are measured in thousand dollars (1000 \$) whereas logged asset and offering amount are used in estimation. 2. Mshare and Fshare are measured in percentage. 3. The crisis dummy is equal to 1 if year is in between 2007 and 2010.

4.3 Correlation Analysis

To motivate our selected control covariates, we start by presenting some simple correlation analysis between the cumulative rating outcomes and control variables in Table 3. If the control varibles could indeed capture the effect of CRA-issuer liasion on ratings, we ought to see some co-movement between them: issuers that are close to Moody's ownership-wise should be assigned higher ratings. In Table 3, we divide the whole sample into the before and after Dodd-Frank periods, recognizing the structural change of regulatory environment.

For Mshare, the correlations are positive and consistent, suggesting that a strong firm-Moody tie always corresponds to a upward pressure for ratings, though the magnitudes overall decease after the Dodd-Frank. But for Fshare, negative correlation is found for investment-level grades, especially after the regulatory change. The magnitudes of having at least one influential common investor are relatively small to the other measures and its effects are even lower after the reform. To sum up, our control covariates do convey some predictive power on ratings and should not be simply left out of the model. From the correlation table, we can also conjecture that the effect of CRA-issuer relation on ratings might be highly heterogeneous. For bonds with extremely high or low ratings, the correlation between the rating and the control variables is quite small. And our empirical model is able to capture such heterogeneity as the functional form of the soft adjustment is modeled flexibly other than the single-index restriction. Besides, it suggests the importance of the financial reform in 2010 and hence demands separate estimation of the structural bond rating models, which is already in our considerations.

5 Empirical Results

We report results in four subsections. First, we estimate a series of parametric rating models in the literature, and argue that the parameters of interest are sensitive to model specification. Hence, it is important to consider a more flexible approach to understand the rating process. In the second and third subsection, we estimate the structural rating model

Table 3: Correlation between Control Variables and Rating Outcomes

	Before DF: 2000-2010 (N=6,618)								
	Aaa	>=Aa	>=A	>=Baa	>=Ba	>=B	Mshare	Fshare	largeSH
Mshare	0.106	0.430	0.498	0.455	0.392	0.130	1.000		
Fshare	-0.018	0.073	0.101	0.097	0.125	0.082	0.470	1.000	
largeSH	0.047	0.220	0.150	0.112	0.079	0.026	0.463	0.191	1.000
	After DF: 2011-2016 ($N = 4,516$)								
	Aaa	>=Aa	>=A	>=Baa	>=Ba	>=B	Mshare	Fshare	largeSH
Mshare	0.166	0.216	0.375	0.342	0.308	0.165	1.000		
Fshare	-0.018	-0.166	-0.034	0.048	0.107	0.128	0.556	1.000	
largeSH	0.045	-0.062	0.044	0.001	0.016	-0.010	0.587	0.560	1.000

Note: 1. >=X represents cumulative ratings above or equal to notch X.

proposed in Section 2 and reports the estimates of index parameters and soft adjustment. Lastly, we discuss patterns of threshold parameters over a changing shareholding relationship.

5.1 Parametric Results

The results of parametric regressions are presented in Table 4. The first three specifications are estimated using ordered probit models while the last two are from ordered logit models. Recall that a CRISIS dummy with full interactions with all covariates are deployed to control for the financial crisis effect for the period before Dodd-Frank.

There are two main findings. First, some regression coefficients change significantly after the passage of Dodd-Frank. PROFIT and OFFAMT have a much larger impact on default risk in the post Dodd-Frank period, indicating that Moody's has become more stringent on firms with low profitability ratio and high debt. The disparity reconfirms the existence of a structural break of rating models. Second, we note that the regression coefficients are sensitive to the distributional assumption of the error term. Oprobit-2 and Ologit-1 (or Oprobt-3 and Ologit-2) differs only by the error distribution. Despite most of the regression coefficients have the same signs as predicted in both specifications, the logit regression coefficients on firm characteristics (ASSET, CVTA, LEVERAGE and PROFIT) are nearly twice as large as the probit coefficients.[26] The somewhat inconsistency results from different

[26]Most coefficients have the correct predicted signs across the board: the amount of total asset, profitability

parametric specifications calls for a robust approach which does not overly restrict the error distribution.

5.2 Semiparametric First Stage: Index Parameters

In this section, we estimate the normalized index parameters defined in Section 2. Recall that these index parameters reflect how ratings are driven by observed "hard information", such as ASSET, LEVERAGE, etc. To fix idea, we focus on comparing the estimation results along two dimensions: (i) between the semiparametric framework proposed in this paper and the baseline parametric model, and (ii) before and after the passage of Dodd-Frank.

The estimation results are shown in Table 5 and 6, respectively for the period before and after Dodd-Frank. To facilitate comparison, the first column of each table "Oprobit-R" gives ratios of estimated coefficients relative to ASSET and Mshare, the variables that we choose to normalize on, respectively.[27] Estimation results of the suggested semiparametric double-index model, "Semi-R", is reported in the last column. As opposed to column 1, we allow arbitrary interactions between shareholding relation with other characteristics and do not need to specify the error term distribution. Another interesting experiment we do, in order to assess the impact of ignoring soft adjustment, is estimating a single-index model "Semi-X" in which we deliberately drop the shareholding relation index $R_i(\alpha)$.

We find the shareholding relation and the CRISIS dummy have very different estimated impacts when switching from ordered-probit to the semiparametric framework. Taking the number of common large shareholder (largeSH) as an example. In the post Dodd-Frank period (reported in Table 6), Oprobit-R predicts a significantly negative impact on default risk, suggesting that Moody's ratings are more favorable to issuers who are invested by

and being a senior bond all have negative impact on the default risk index and thus leads to higher ratings. On the other hand, high variance of assets, measured by CVTA as well as high leverage ratios are in accordance with a larger default risk index.

[27] The standard errors of parameter ratios are calculated using the delta-method: via a first order Taylor expansion around true values, $\widehat{\beta}_2/\widehat{\beta}_1 - \beta_{20}/\beta_{10} \approx -(\widehat{\beta}_1 - \beta_{10})\beta_{20}/\beta_{10}^2 + (\widehat{\beta}_2 - \beta_{20})/\widehat{\beta}_1$ and then compute the standard error. Admittedly, the semiparametric model can identify only the relative ratios instead of each single coefficients. To construct comparable default risk index, one can back out individual coefficients assuming $\widehat{\beta}_1$ is close to that of ordered probit model.

Table 4: Parametric Specifications of Preliminary Estimates

	Before DF: 2000-2010					After DF: 2011-2016				
Variables	Oprobit-1	Oprobit-2	Oprobit-3	Ologit-1	Ologit-2	Oprobit-1	Oprobit-2	Oprobit-3	Ologit-1	Ologit-2
ASSET	-0.587	-0.577	-0.485	-1.131	-0.980	-0.506	-0.500	-0.432	-0.911	-0.792
	(0.010)	(0.012)	(0.014)	(0.023)	(0.028)	(0.013)	(0.013)	(0.014)	(0.025)	(0.027)
CVTA	0.394	0.554	0.634	1.335	1.399	0.571	0.552	0.618	0.990	1.092
	(0.099)	(0.123)	(0.124)	(0.223)	(0.225)	(0.106)	(0.105)	(0.107)	(0.187)	(0.191)
LEVERAGE	2.336	2.326	2.289	4.096	4.090	2.054	2.074	2.099	3.845	3.899
	(0.103)	(0.123)	(0.123)	(0.230)	(0.233)	(0.129)	(0.128)	(0.129)	(0.230)	(0.232)
PROFIT	-9.649	-9.192	-8.457	-21.376	-19.406	-16.620	-16.488	-15.265	-29.883	-27.718
	(0.278)	(0.332)	(0.341)	(0.781)	(0.799)	(0.386)	(0.384)	(0.398)	(0.736)	(0.750)
OFFAMT	0.018	0.033	0.057	0.056	0.091	0.179	0.174	0.188	0.325	0.352
	(0.009)	(0.013)	(0.013)	(0.022)	(0.022)	(0.025)	(0.025)	(0.025)	(0.044)	(0.045)
SENIOR	-0.523	-0.572	-0.538	-1.056	-1.021	-0.817	-0.806	-0.784	-1.376	-1.351
	(0.038)	(0.045)	(0.046)	(0.082)	(0.083)	(0.062)	(0.062)	(0.062)	(0.117)	(0.118)
Mshare			-0.026		-0.040			-0.025		-0.046
			(0.002)		(0.004)			(0.002)		(0.003)
Fshare			0.011		0.018			0.007		0.012
			(0.001)		(0.002)			(0.001)		(0.002)
largeSH			0.081		0.099			0.654		1.245
			(0.039)		(0.070)			(0.077)		(0.140)
CRISIS		0.634	1.354	1.408	2.734					
		(0.326)	(0.353)	(0.588)	(0.645)					
CRISIS*VAR		Y	Y	Y	Y					
Year	Y					Y				
Obs. N	6618	6618	6618	6618	6618	4516	4516	4516	4516	4516

Note: 1. CRISIS*VAR indicates the inclusion of full interaction terms with the crisis dummy. 2. Oprobit-1 is the specification controlling for year fixed effects. 3. s.e. are presented in parenthesis.

its own large shareholders. In contrast, Semi-R predicts a positive impact with similar magnitude, suggesting Moody's has become more stringent on these related firms. The disparity is not only statistically significant, but economically large. Estimated impact of the CRISIS dummy also differs across the board. From Semi-R, most of the interaction terms have insignificant coefficients, meaning that firm and bond characteristics roughly have the same impact on ratings in and out of economic downturns. In contrast, from Oprobit-R, we find firm's financial stability (CVTA) and profitability (PROFIT) have significantly less impacts on rating during the crisis period. As can be seen in Semi-X, interaction terms also have differential impacts on ratings after we drop the shareholding relation variables.

Turning to the comparison before and after the Dodd-Frank, the differences in estimated coefficients are equally striking. First, Moody's attention to firm and bond characteristics has clearly changed over time. Specifically, the impact of firm stability (CVTA) decreases by half after the passage of Dodd-Frank, whereas the impact of profitability (PROFIT) increases by half. The biggest difference comes from the relative importance of issuing amount: the estimated effect increasing by a factor of ten, reflecting Moody's has increased its scrutiny on the amount of debt that a issuer is taking. Second, the impact of shareholding relation also changes over time. In particular, relationship with Moody's large shareholders led to higher ratings before the Dodd-Frank. Such effects, however, reverse sign and enlarge in magnitude after the Dodd-Frank.

The aforementioned findings suggest that Moody's rating model has changed significantly after the Dodd-Frank, in terms of its relative focus on specific characteristics as well as its treatment to issuers in terms of shareholding relations. Recall that we use the shareholding relation to "anchor" the amount of soft information that the CRA may receive from common shareholders. Thereby, the differential impact of shareholding relation before and after the Dodd-Frank may reflect a substantial change in terms of how the CRA utilize soft information to determine ratings. To investigate this issue further, in the next section we estimate the soft adjustment for each bond issuer given its shareholding relation with Moody's and report the distributional pattern of soft adjustment.

Table 5: Estimation Results of Creditworthiness Index Parameters before Dodd-Frank Act

	Parametric	Semiparametric	
Variables	Oprobit-R	Semi-X	Semi-R
Structural Financial Risk Parameters			
CVTA	-1.306 ***	-1.773 ***	-1.884 ***
	(0.259)	(0.204)	(0.213)
LEVERAGE	-4.715 ***	-3.467 ***	-3.862 ***
	(0.320)	(0.254)	(0.298)
PROFIT	17.422 ***	24.339 ***	23.919 ***
	(0.706)	(0.697)	(0.788)
OFFAMT	-0.118 ***	-0.014	-0.047 **
	(0.027)	(0.016)	(0.019)
SENIOR	1.109 ***	0.928 ***	0.994 ***
	(0.101)	(0.079)	(0.082)
CRISIS	-2.790 ***	-1.794 ***	-1.702 ***
	(0.757)	(0.514)	(0.575)
CRISIS*ASSET	0.222 ***	0.110 ***	0.168 ***
	(0.045)	(0.029)	(0.034)
CRISIS*CVTA	0.821 **	0.310	0.073
	(0.401)	(0.314)	(0.329)
CRISIS*LEVERAGE	0.117	0.898 **	0.440
	(0.444)	(0.366)	(0.443)
CRISIS*PROFIT	3.850 ***	-1.994 *	-1.359
	(1.184)	(1.148)	(1.155)
CRISIS*OFFAMT	0.044	0.031	-0.011
	(0.038)	(0.022)	(0.027)
CRISIS*SENIOR	-0.389 **	-0.575 ***	-0.526 ***
	(0.174)	(0.148)	(0.173)
Control Index Parameters			
Fshare	-0.444 ***		-1.288 ***
	(0.106)		(0.036)
largeSH	-3.185 *		-1.903 **
	(1.735)		(0.812)
CRISIS*Mshare	-0.765 ***		0.046
	(0.247)		(0.048)
CRISIS*Fshare	0.428 ***		-0.068
	(0.120)		(0.049)
CRISIS*largeSH	1.558		-6.370 ***
	(2.760)		(1.423)
ASSET	-0.485 ***		
Mshare	-0.026 ***		
$N = 6618$			

Note: 1. The estimation uses data from 2000 to 2010. 2. Estimates represent normalized coefficient ratios with respect to log of asset and Mshare, respectively for financial and control parameters. 3. Oprobit-R is estimated by MLE. Semi-X and semi-R are estimated by pseudo-MLE. 4. The rule-of-thumb bandwidths, $h = 1.06 s.d.(R) N^{-r}$ are used, with the optimal rate i.e. $r = 1/6$ for double index models. 5. Standard errors are in parentheses. 6. Significant level: *10 percent, **5 percent, ***1 percent.

Table 6: Estimation Results of Creditworthiness Index Parameters after Dodd-Frank Act

	Parametric	Semiparametric	
Variables	Oprobit-R	Semi-X	Semi-R
Structural Financial Risk Parameters			
CVTA	-1.397 ***	-0.447 ***	-0.915 ***
	(0.265)	(0.182)	(0.263)
LEVERAGE	-4.893 ***	-3.140 ***	-3.939 ***
	(0.380)	(0.282)	(0.343)
PROFIT	35.361 ***	28.975 ***	32.416 ***
	(1.191)	(0.890)	(1.416)
OFFAMT	-0.445 ***	-0.558 ***	-0.430 ***
	(0.055)	(0.041)	(0.065)
SENIOR	1.834 ***	1.308 ***	1.154 ***
	(0.148)	(0.158)	(0.215)
Control Index Parameters			
Fshare	-0.279 ***		-2.409 ***
	(0.091)		(0.181)
largeSH	-26.150 ***		26.308 ***
	(4.886)		(4.773)
ASSET	-0.430 ***		
Mshare	-0.025 ***		
$N = 4516$			

Note: 1. The estimation uses data from 2011 to 2016. 2. Estimates represent normalized coefficient ratios with respect to log of asset and Mshare, respectively for financial and control parameters. 3. Oprobit-R is estimated by MLE. Semi-X and semi-R are estimated by pseudo-MLE. 4. The rule-of-thumb bandwidths, $h = 1.06s.d.(R)N^{-r}$ are used, with the optimal rate i.e. $r = 1/6$ for double index models. 5. Standard errors are in parentheses. 6. Significance level: *10 percent, **5 percent, ***1 percent.

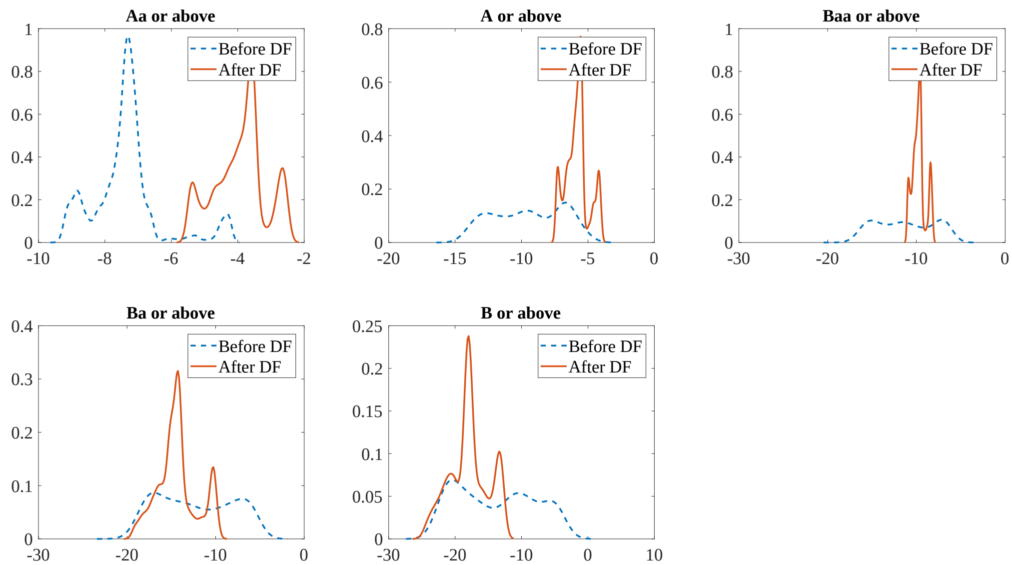
5.3 Second Stage: Soft Adjustment

In this section, we estimate Moody's soft adjustment to each bond issuers: that, say, to what extent would Moody's alter the threshold parameters T_{ji} given a issuer's shareholding relation with Moody's. To fix idea, recall that the (unobserved) soft adjustment is represented by the relative conditional mean thresholds $\delta_{j,0}(r) \equiv E[T_{ji} - T_{0i}|R_i(\alpha_0) = r]$. To have a comparable measure of soft adjustment, we choose to normalize the baseline level, Aaa notch that is unlikely to be affected by the shareholding relation. In the appendix, we provide some practical and empirical support for choosing this base category.

Figure 3 depicts the empirical distributions of estimated individual-bond soft adjustment $\Delta_{j,0}(r)$ from the semiparametric model, with the dash (dotted) line indicating the period before(after) the Dodd-Frank. We only plot the distribution for the five ratings categories from Aa to B, as the adjustment for Aaa category is already normalized to zero.

We focus on two implications of the results. First, the dispersion of soft adjustment decreases substantially after the Dodd-Frank as the empirical distribution becomes more concentrated around the mean, especially for bonds with median level of credit worthiness. Since soft adjustment, in some sense, reflects the CRA's private information, the decline of soft adjustment is consistent with the objective of the Dodd-Frank Act to increase the transparency of rating methodologies. Second, the distribution of soft adjustment shift in means after the Dodd-Frank. For investment grade bonds, the soft adjustment, on average, shift towards the right, implying more stringent rating criteria. Put it differently, receiving investment grade has become more difficult after 2010 for observationally identical bonds. For speculative bonds of Ba or below, the mean thresholds have become smaller, indicating more relaxed criteria of the CRA. Besides, our plots of densities exhibit the apparent tri-modal feature, especially for the investment grade bonds. Again, estimation of the distribution of soft adjustment has been difficult in the parametric context. Strict distributional assumptions, usually a normal random variable with unknown means, have to be imposed to permit estimation. However, as can be seen from Figure 3 in which the soft adjustments are estimated in a distribution-free manner, the normality assumption on soft adjustments is highly suspicious.

Figure 3: Empirical Distributions of Soft Adjustment before and after the Dodd-Frank Act



Note: 1. Sample periods. Before the Dodd-Frank: 2000-2010; After the Dodd-Frank: 2011-2016. 2. The soft adjustment is estimated as mean thresholds relative to the base level Aaa using the two-step semiparametric estimator. 4. The rule-of-thumb bandwidths, $h = 1.06s.d.(R)N^{-r}$ are used, with the optimal rate i.e. $r = 1/6$ for double index models.

Recall that the proposed semiparametric model allows $\delta_{j,0}$ to be driven by $R_i(\alpha_0)$, a bond issuer's relationship with Moody's, whereas the ordered-probit assumes that $\delta_{j,0}$ is a constant. In Table 7, we report estimation results from the two approaches to highlight the heterogeneity in threshold captured by the semiparametric model. The first two columns report relative thresholds from ordered probit/logit specifications,[28] and the third column from the single-index semiparametric model without assuming the error term distribution.[29] Noticeable differences between Semi-X and the two parametric models suggests that neither ordered probit nor logit correctly describes the underlying data. Turning to the proposed model Semi-R, we present estimated thresholds conditional on various percentiles of the control index. It can be inferred from the heterogeneous pattern of threshold that the extent of soft adjustment varies with the control index $R_i\alpha_0$. Comparing the standard errors of

[28]The base level is Aaa. Relative thresholds are defined as $(T_{ji} - T_{0i})/\widehat{\beta}_1$, where $\widehat{\beta}_1$ is the estimated log asset coefficient.

[29]The bootstrapped standard errors are presented in parentheses for the semiparametric models and we use the delta-method to compute them for Ologit and Oprobit similar to those of relative coefficients.

$\delta_{j,0}$ before and after the Dodd-Frank period, the smaller standard errors in the later period confirms the earlier finding that soft adjustment become less dispersed.

5.4 Patterns of Soft adjustment over Shareholder Relationship

Following our discussion on the threshold parameters $\Delta_{j,0}$, in this section we provide in-depth analysis on how exactly are soft adjustments driven by the shareholding relation. Recall that the shareholding relation $R_i(\alpha_0) \equiv \text{Mshare}_i + \alpha_1 \text{Fshare}_i + \alpha_2 \text{LargeSH}_i$. We assess the pattern of $\Delta_{j,0}(r)$ over both $R_i(\alpha_0)$ and the three determinants, having in mind that $\alpha_1 < 0$ and $\alpha_2 < 0$ before the Dodd-Frank and $\alpha_1 < 0$ and $\alpha_2 > 0$ after the Dodd-Frank.

In Figure 4, we plot the relationship between $R_i(\alpha_0)$ and the soft adjustment across different rating categories, with the left(right) panel indicating the period before(after) the Dodd-Frank. First, soft adjustment and $R_i(\alpha_0)$ poses a "U-shape" relationship before the Dodd-Frank. As the bond issuer builds a tighter relationship with common shareholders (a higher Fshare and/or LargeSH inducing a lower $R_i(\alpha_0)$) Moody's starts to relax its rating criteria: it is easier for firms with a stronger connection to receive higher ratings. Possibly due to the worry of conflicts of interest, Moody's starts to tighten its rating criteria when relationship gets too strong. This pattern is nearly uniform for all rating categories before the Dodd-Frank, with Aa being the only exception. Interestingly, for the period after the Dodd-Frank, such pattern has changed completely: as $R_i(\alpha_0)$ strengthens, Moody's uniformly tightens the rating criteria.[30] Given that largeSH has a positive impact of $R_i(\alpha_0)$ in this period(e.g. $\alpha_2 > 0$), we conclude that Moody's has become more stringent on issuers related with its large shareholders after the Dodd-Frank. This change of pattern could be a result of rating agency's greater concern of conflicts of interest.

Another sharp contrast between the two panels in Figure 4 is threshold parameters for different rating categories converge to each other before the Dodd-Frank, but not after. Recall that those parameters partition the latent default risk into different rating categories.

[30]Since in the post Dodd-Frank period, largeSH and Fshare have opposite impact on $R_i(\alpha_0)$, it is unclear whether a increasing $R_i(\alpha_0)$ represents a stronger shareholding relation. Therefore we provide more analysis on the impact of individual variable on soft adjustment later.

Table 7: Estimation Results of Soft Adjustment ($\widehat{\delta}$) at Control Index Percentiles

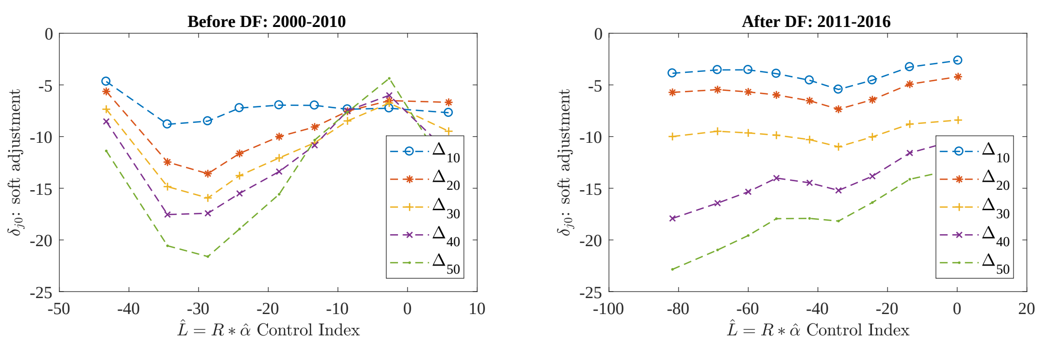
	Oprobit	Ologit	Semi-X	Semi-R								
				.1	.2	.3	.4	.5	.6	.7	.8	.9
Case 1: 2000-2010 Before Dodd-Frank												
$\widehat{\delta}_{0,1}(R)$	-4.281	-4.137	-6.863	-4.674	-8.813	-8.504	-7.236	-6.958	-6.982	-7.355	-7.258	-7.691
	(0.323)	(0.322)	(1.333)	(1.135)	(1.232)	(1.198)	(0.999)	(0.977)	(1.032)	(1.209)	(1.423)	(1.620)
$\widehat{\delta}_{0,2}(R)$	-7.453	-7.053	-8.340	-5.629	-12.466	-13.588	-11.632	-10.004	-9.073	-7.545	-6.520	-6.686
	(0.431)	(0.416)	(1.407)	(1.298)	(1.712)	(2.131)	(2.283)	(1.671)	(1.467)	(1.519)	(1.550)	(2.187)
$\widehat{\delta}_{0,3}(R)$	-10.766	-10.145	-10.978	-7.352	-14.842	-15.938	-13.779	-12.064	-10.601	-8.485	-6.754	-9.483
	(0.535)	(0.507)	(1.607)	(1.498)	(1.852)	(2.362)	(2.418)	(1.921)	(1.959)	(2.028)	(2.310)	(2.980)
$\widehat{\delta}_{0,4}(R)$	-12.601	-11.860	-12.411	-8.536	-17.543	-17.422	-15.502	-13.386	-10.845	-7.498	-6.007	-11.544
	(0.592)	(0.557)	(1.793)	(2.003)	(2.176)	(2.593)	(2.560)	(2.422)	(2.972)	(3.293)	(4.044)	(3.256)
$\widehat{\delta}_{0,5}(R)$	-16.676	-16.082	-14.627	-11.387	-20.567	-21.608	-18.938	-15.581	-10.341	-7.626	-4.369	-13.453
	(0.719)	(0.681)	(2.471)	(4.462)	(4.273)	(4.283)	(3.827)	(3.885)	(4.540)	(4.460)	(5.640)	(3.489)
Case 2: 2011-2016 After Dodd-Frank												
$\widehat{\delta}_{0,1}(R)$	-2.667	-2.734	-3.676	-3.874	-3.550	-3.541	-3.909	-4.552	-5.441	-4.544	-3.268	-2.637
	(0.231)	(0.252)	(0.856)	(0.277)	(0.272)	(0.327)	(0.421)	(0.711)	(0.577)	(0.414)	(0.493)	(0.650)
$\widehat{\delta}_{0,2}(R)$	-6.372	-6.334	-5.333	-5.723	-5.460	-5.678	-5.968	-6.529	-7.355	-6.433	-4.929	-4.204
	(0.351)	(0.371)	(0.861)	(0.267)	(0.277)	(0.401)	(0.526)	(0.775)	(0.615)	(0.425)	(0.514)	(0.724)
$\widehat{\delta}_{0,3}(R)$	-11.165	-11.047	-8.882	-9.992	-9.486	-9.645	-9.866	-10.294	-10.987	-10.027	-8.790	-8.402
	(0.496)	(0.514)	(0.879)	(0.317)	(0.319)	(0.453)	(0.550)	(0.802)	(0.674)	(0.478)	(0.557)	(0.762)
$\widehat{\delta}_{0,4}(R)$	-13.037	-12.888	-12.339	-17.919	-16.441	-15.341	-14.015	-14.473	-15.208	-13.837	-11.586	-10.280
	(0.552)	(0.569)	(0.901)	(0.609)	(0.773)	(0.849)	(0.681)	(0.893)	(0.729)	(0.538)	(0.612)	(0.773)
$\widehat{\delta}_{0,5}(R)$	-17.775	-17.817	-15.064	-22.847	-20.972	-19.572	-17.944	-17.917	-18.174	-16.381	-14.128	-13.094
	(0.700)	(0.726)	(0.878)	(0.587)	(0.751)	(0.793)	(0.597)	(0.837)	(0.703)	(0.499)	(0.604)	(0.780)

Note: 1. Each soft adjustment $\widehat{\delta}(\cdot)$ is evaluated at the percentile of R or $\widehat{L}$. For instance, .1 denotes 10 percentile. 2. Scaled thresholds estimates are computed relative to the base level of $Y = 0$, indicating the Aaa notch. 3. Semi-X refers to semiparametric estimation without controlling soft information and Semi-R does the control. 4. In parenthesis are the standard errors. Those of Oprobit and Ologit are computed using first order Delta approximation. Those of semiparametric models are bootstrapped s.e.(s) with 50 times of draws.

It can be seen from the left panel that when the relationship index takes value between -10 and 0 (about 70-80 percentile), threshold parameters for different rating categories become indistinguishable. In terms of rating criteria, this indicates Moody's does not have a clear criteria to separate the safe bonds from the bad bonds issued by highly connected firms. Instead, actual rating assignments must involve discretion. This pattern also disappear after the Dodd-Frank, as the threshold parameter maintain an ordered relationship throughout.

The contrast of patterns suggests the Dodd-Frank has improved the transparency of credit rating procedures. In particular, we have witnessed that Moody's has a clearer picture on mapping latent default risk into different rating categories. In addition, Moody's has also become more stringent when rating bonds issued by closely connected firms, as the threshold parameters increase monotonically as the relationship index strengthens. In the online appendices, we further examine the empirical relationship between the soft adjustment and each of the shareholding measures: Mshare, Fshare and LargeSH, for various counterfactual scenarios. The patterns of $\Delta_{j,0}$ over all three measures are roughly the same, albeit to minor disparity, as if we only use the aggregate the measure $R_i(\alpha_0)$. Detailed discussion can be found in the online appendices.

Figure 4: Estimated Relationship between Soft Adjustment and Shareholding Control Index



Note: 1. Left panel: before the Dodd-Frank; right panel: after the Dodd-Frank. 2. Y-axis plots the estimated soft adjustment as conditional mean thresholds relative to Aaa level. 3. X-axis plots various percentiles of estimated control index for shareholding relations.

6 Conclusions

This paper considers the role of soft adjustment when estimating a structural corporate bond rating model used by CRAs and empirically assesses the extend of reduction of the conflict-of-interest after the enactment of the Dodd-Frank Act, specific to the credit rating industry. From an empirical point of view, the presence of hidden soft adjustment could cause endogenous determination of firm and bond characteristics, resulting in inconsistent default risk index estimates and incorrect rating probability functions. To resolve this identification issue, we model the soft adjustment as the bond-specific stochastic thresholds in a fully nonparametric way and suggest to approximate it using the shareholding structures between a bond-issuer and a publicly listed CRA. In terms of econometrics, we recover the unobserved soft adjustment by extending the *shift restriction* to ordered response models with endogenous regressors. The empirical method in this paper can be applied to other contexts featuring ordered response with unobserved heterogeneous thresholds such as subjective health status, life happiness, etc.

For empirics, we focus on initial bond ratings of listed firms after the Moody's went public in 2000 and until 2016, covering a fews year after the passage of the Dodd-Frank Act. Our empirical results suggest that there is a significant reduction of soft adjustment over all rating categories in terms of dispersion after the reform. We also find that it becomes more difficult to be rated as investment grade bonds on average, reflected by the shift of mean thresholds. To shed light on the rationale behind, we find the primary factor that drives the heterogeneous adjustment is how much common shareholders own a particular bond-issuing firm instead of the amount of investment in Moody's.

Our model does not consider the competition effect in the rating process, especially given the "triopoly" market structure of the U.S. credit rating industry. Under the current issuer-pays model, the CRA charge fees to issuers whose debt will be rated by the same agency. This would produce another aspect of conflict-of-interest where CRAs attempt to make more profit and attract more clients by lowering standards. However, our measure of soft adjustment is not designed to capture such bias. Future studies may consider incorporating strategic rating behaviors among CRAs.

Appendix A Identification Proof

Proof of proposition 1. For proposition 1, the proof of shift restriction resembles that in Klein and Sherman (2002) but the only difference is that every step needs to be conditional on R_i. For each $(v, r) \in \mathbb{R} \times \mathcal{R}$ and $v = x'\beta_0$,

$$\begin{aligned}
P_j(v,r) &\equiv \Pr(Y_i \leq j | V_{0i} = v, R_i = r) \\
&= \Pr(v \leq T_{ji} | V_{0i} = v, R_i = r) \\
&= \Pr(v - t_j(r) \leq u | R_i = r) \\
&= \Pr(v + \Delta_{j,k}(r) - t_k(r) \leq u | R_i = r) \\
&= \Pr(v + \Delta_{j,k}(r) \leq T_{ki} | R_i = r) \\
&= \Pr(Y_i \leq k | V_{0i} = v + \Delta_{j,k}(r), R_i = r) \\
&= P_k(v + \Delta_{j,k}(r), r)
\end{aligned}$$

Note that A-I.2 ensures the third to last equalities.

For point identification of $\Delta_{j,k}(r)$, note that $P_k(\cdot, r)$ is strictly increasing and therefore invertible because it is the conditional distribution function of u given $R_i = r$. So $\Delta_{j,k}(r) = P_k^{-1}(P_j(v,r), r) - v$, where $P_k^{-1}(P_k(v,r), r) = v$. Then by construction, $\Delta_{j,k}(r)$ is identified.

The identification of $\Delta_{j,k}^c(\cdot)$ is established as follows,

$$\begin{aligned}
\Delta_{j,k}^c(x) &= E(T_{ki} - T_{ji} | X_i = x) \\
&= E[E(T_{ki} - T_{ji} | X_i = x, R_i) | X_i = x] \\
&= E[\Delta_{j,k}(R_i) | X_i = x]
\end{aligned}$$

The second equality uses the iterative expectation by firstly conditional on both X_i and R_i.

The last equality holds by definition.

□

Appendix B Choice of Base Category

In this section we provide some support for our choice of base category in relation to Assumption A-I.2, in which we assume that Aaa-rated bonds do not contain any heterogeneous soft adjustment that is induced or related to shareholder structures. In other words, it implies that the soft adjustment is constant across Aaa-rated bonds.

From a methodological point of view, Aaa bonds have exceptional creditworthiness and are investment grade bonds with minimum risk. Issuers of those bonds usually have outstanding financial performance and can easily meet issuing commitments. Aaa bonds are also faced with more supervision from financial regulators and investors, for its signaling function of the health of the financial system as a whole. As a result, CRAs' have very strict and prudent criteria to select and assign Aaa ratings and the conflict of interest is unlikely to occur at this level of creditworthiness in fear of reputation loss.

We perform additional parametric regressions to roughly test this proposition. In Table 8, both probit and logit results are shown for before and after the Dodd-Frank, along with the F-statistics that are used to test the joint significance of all three relationship controls. The dependent variable is a dummy indicating whether the bond is rated Aaa or not. We only focus on subsamples of bonds of adjacent level, say Aa, as those bear much closer relationship. The small test statistics indicate that our proposed shareholder controls are not jointly significant at 1% level. For those before the Dodd-Frank, they are not significant at 5% even with a decent amount of observations. Such results cannot be found for other categories. This empirical fact provides additional support to assume away shareholding-induced heterogeneity for Aaa category.

Table 8: Preliminary Regression on the Probability of Being Rated to Aaa Notch

	Before DF		After DF	
Variables	Probit	Logit	Probit	Logit
ASSET	-0.330	-0.751	0.280	0.418
	(0.084)	(0.169)	(0.226)	(0.415)
CVTA	0.257	0.448	-15.316	-26.782
	(0.601)	(1.064)	(2.602)	(4.596)
LEVERAGE	-4.773	-11.875	-22.959	-40.870
	(1.385)	(3.081)	(3.116)	(5.891)
PROFIT	7.644	14.131	27.412	46.894
	(2.298)	(4.329)	(5.357)	(9.369)
OFFAMT	0.103	0.160	0.286	0.456
	(0.060)	(0.126)	(0.164)	(0.285)
SENIOR	-0.425	-1.087		
	(0.334)	(0.707)		
Mshare	-0.001	0.001	0.004	0.010
	(0.008)	(0.015)	(0.013)	(0.021)
Fshare	-0.013	-0.031	0.061	0.105
	(0.007)	(0.015)	(0.030)	(0.051)
largeSH	0.185	0.424		
	(0.277)	(0.548)		
F-stat	5.110	6.290	5.740	6.230
P-value	0.164	0.098	0.057	0.044
Obs. N	1092	1092	290	290

Note: 1. Subsample of Aaa or Aa is used instead of full sample. 2. F-statistics are used to test the joint significance of Mshare, Fshare and largeSH, with p-values below. 3. SENIOR and largeSH are omitted for those after DF due to lack of variation. 4. s.e. are presented in parenthesis.

References

Hyungtaik Ahn, Hidehiko Ichimura, and James L Powell. Simple estimators for monotone index models. 1996.

Christopher Alessi, Roya Wolverson, and Mohammed Aly Sergie. The credit rating controversy. *Council on Foreign Relations*, 2013.

Bo Becker and Todd Milbourn. How did increased competition affect credit ratings? *Journal of Financial Economics*, 101(3):493–514, 2011. doi: 10.1016/j.jfineco.2011.03.012.

Marshall E Blume, Felix Lim, and A Craig MacKinlay. The declining credit quality of us corporate debt: Myth or reality? *The Journal of Finance*, 53(4):1389–1413, 1998. doi: 10.1111/0022-1082.00057.

Richard W Blundell and James L Powell. Endogeneity in semiparametric binary response models. *The Review of Economic Studies*, 71(3):655–679, 2004. doi: 10.1111/j.1467-937X.2004.00299.x.

John C Bogle. *The battle for the soul of capitalism.* Yale University Press, 2005.

John Y Campbell and Glen B Taksler. Equity volatility and corporate bond yields. *The Journal of Finance*, 58(6):2321–2350, 2003. doi: 10.1046/j.1540-6261.2003.00607.x.

Richard Cantor, Frank Packer, et al. The credit rating industry. *Quarterly Review*, (Sum): 1–26, 1994.

Songnian Chen. Semiparametric estimation of a location parameter in the binary choice model. *Econometric Theory*, 15(01):79–98, 1999.

Songnian Chen. Rank estimation of a location parameter in the binary choice model. *Journal of Econometrics*, 98(2):317–334, 2000. doi: 10.1016/S0304-4076(00)00021-X.

Jess Cornaggia and Kimberly J Cornaggia. Estimating the costs of issuer-paid credit ratings. *Review of Financial Studies*, 26(9):2229–2269, 2013. doi: 10.1093/rfs/hht041.

Jean-Pierre Florens, James J Heckman, Costas Meghir, and Edward Vytlacil. Identification of treatment effects using control functions in models with continuous, endogenous treatment and heterogeneous effects. *Econometrica*, 76(5):1191–1206, 2008. doi: 10.3982/ECTA5317.

Stefan Hoderlein and Robert Sherman. Identification and estimation in a correlated random coefficients binary response model. *Journal of Econometrics*, 2015. doi: 10.1016/j.jeconom.2015.03.044.

Joel L Horowitz. A smoothed maximum score estimator for the binary response model. *Econometrica: journal of the Econometric Society*, pages 505–531, 1992.

Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun-Hwa Chen, and Soushan Wu. Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision support systems*, 37(4):543–558, 2004. doi: 10.1016/S0167-9236(03)00086-1.

Hidehiko Ichimura and Lung-Fei Lee. Semiparametric least squares estimation of multiple index models: single equation estimation. In *Nonparametric and semiparametric methods in econometrics and statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics. Cambridge*, pages 3–49, 1991.

Guido W Imbens. Nonadditive models with endogenous regressors. *Econometric Society Monographs*, 43:17, 2007.

Guido W Imbens and Whitney K Newey. Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica*, 77(5):1481–1512, 2009. doi: 10.3982/ECTA7108.

John Xuefeng Jiang, Mary Harris Stanford, and Yuan Xie. Does it matter who pays for bond ratings? historical evidence. *Journal of Financial Economics*, 105(3):607–621, 2012. doi: 10.1016/j.jfineco.2012.04.001.

Yixiao Jiang. Did shareholders affect the credit rating decision? evidence from a semiparametric ordered model with marginal effects analysis. Technical report, Rutgers University, Department of Economics, 2017. Working paper.

Robert S Kaplan and Gabriel Urwitz. Statistical models of bond ratings: A methodological inquiry. *Journal of Business*, pages 231–261, 1979.

Maximilian Kasy. Identification in triangular systems using control functions. *Econometric Theory*, 27(03):663–671, 2011. doi: 10.1017/S0266466610000460.

Simi Kedia, Shivaram Rajgopal, and Xing Zhou. Did going public impair moody s credit ratings? *Journal of Financial Economics*, 114(2):293–315, 2014. doi: 10.1016/j.jfineco.2014.07.005.

Simi Kedia, Shivaram Rajgopal, and Xing Zhou. Does it matter who owns moody's, 2016.

Roger Klein and Chan Shen. Bias corrections in testing and estimating semiparametric, single index models. *Econometric Theory*, 26(06):1683–1718, 2010. doi: 10.1017/S0266466609990764.

Roger Klein and Francis Vella. A semiparametric model for binary response and continuous outcomes under index heteroscedasticity. *Journal of Applied Econometrics*, 24(5):735–762, 2009. doi: 10.1002/jae.1064.

Roger W Klein and Robert P Sherman. Shift restrictions and semiparametric estimation in ordered response models. *Econometrica*, 70(2):663–691, 2002. doi: 10.1111/1468-0262.00299.

Roger W Klein and Richard H Spady. An efficient semiparametric estimator for binary response models. *Econometrica: Journal of the Econometric Society*, pages 387–421, 1993. doi: 10.2307/2951556.

Pepa Kraft. Rating agency adjustments to gaap financial statements and their effect on ratings and credit spreads. *The Accounting Review*, 90(2):641–674, 2014. doi: 10.2308/accr-50858.

Arthur Lewbel. Semiparametric estimation of location and other discrete choice moments. *Econometric Theory*, 13(01):32–51, 1997. doi: 10.1017/S0266466600005636.

Arthur Lewbel. Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97(1):145–177, 2000. doi: 10.1016/S0304-4076(00)00015-4.

Arthur Lewbel. Ordered response threshold estimation. *Unpublished working paper, Boston College*, 2003.

Charles F Manski. Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of econometrics*, 27(3):313–333, 1985. doi: 10.1016/0304-4076(85)90009-0.

Matthew Masten and Alexander Torgovitsky. Instrumental variables estimation of a generalized correlated random coefficients model. Technical report, Centre for Microdata Methods and Practice, Institute for Fiscal Studies, 2014.

Jürgen Maurer, Roger Klein, and Francis Vella. Subjective health assessments and active labor market participation of older men: evidence from a semiparametric binary choice model with nonadditive correlated individual-specific effects. *Review of Economics and Statistics*, 93(3):764–774, 2011. doi: 10.1162/REST_a_00097.

Frank Partnoy. Rethinking regulation of credit rating agencies: An institutional investor perspective. *Council of Institutional Investors, April*, pages 09–014, 2009.

Mitchell A Petersen. Information: Hard and soft. 2004.

George E Pinches and Kent A Mingo. A multivariate analysis of industrial bond ratings. *The Journal of Finance*, 28(1):1–18, 1973. doi: 10.1111/j.1540-6261.1973.tb01341.x.

James L Powell, James H Stock, and Thomas M Stoker. Semiparametric estimation of index coefficients. *Econometrica: Journal of the Econometric Society*, pages 1403–1430, 1989. doi: 10.2307/1913713.

Chan Shen and Roger Klein. Market recursive differencing: Bias reduction with regular kernels. Technical report, Rutgers University, Department of Economics, 2017. Working paper.

Lawrence J White. The credit rating industry: An industrial organization analysis. In *Ratings, rating agencies and the global financial system*, pages 41–63. Springer, 2002. doi: 10.1007/978-1-4615-0999-8_3.