International Association For Quantitative Finance

Welcome

5:45 pm
Registration

6:00 pm
Program

7:30 pm
Reception

For more information on the IAQF visit www.iaqf.org

IAQF & Thalesians Seminar Series:

Semiparametric Estimation of a Credit Rating Model

A Talk by
Yixiao (Ethan) Jiang

Semiparametric Estimation of a Credit Rating Model

Thalesians/IAQF seminar

Yixiao Jiang (Ethan)

Rutgers University

February 12, 2019

Credit Ratings and Conflicts of Interest

- Large literature on credit rating agency (CRA) conflicts of interest and rating quality

Credit Ratings and Conflicts of Interest

- Large literature on credit rating agency (CRA) conflicts of interest and rating quality
- One of the conflicts: shared-ownership
  - CRAs are owned by institutional investors. Ex: Moody's
  - 90% of bonds rated by Moody's in 2001-2016 are issued by firms which are invested by Moody's shareholders
  - First studied in Kedia et al. (2016): find bias towards large shareholders

Credit Ratings and Conflicts of Interest

- Large literature on credit rating agency (CRA) conflicts of interest and rating quality
- One of the conflicts: shared-ownership
  - CRAs are owned by institutional investors. Ex: Moody's
  - 90% of bonds rated by Moody's in 2001-2016 are issued by firms which are invested by Moody's shareholders
  - First studied in Kedia et al. (2016): find bias towards large shareholders
- **This paper** examines the empirical relationship between rating inflation and cross-ownership.
  - leverage a large panel data with more than **1500** Moody's shareholders
  - novel econometric framework

Econometric Motivation

Given a proxy Z for the cross-ownership between issuer i and the CRA, run: $Y_i = f(Z_i, control, \epsilon_i) = \beta_0 Z_i + \pi_0 controls + \epsilon_i$

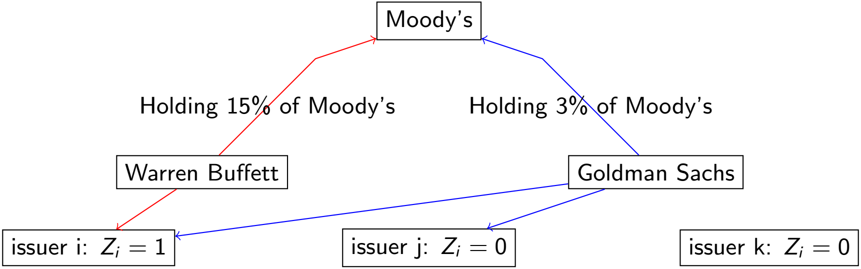
Econometric Motivation

Given a proxy Z for the cross-ownership between issuer i and the CRA, run: $Y_i = f(Z_i, control, \epsilon_i) = \beta_0 Z_i + \pi_0 controls + \epsilon_i$

Concern 1 Z_i **is difficult to find...**

- Kedia et al. (2016) employ a dummy variable approach: whether related to Berkshire Hathaway or not
- Omitting other shareholders is likely to bias $\hat{\beta}$ upward.

Example: omitted variable bias



Under the regression: $Y_i = \beta_0 Z_i + \pi_0 controls + \epsilon_i$

- β_0 captures the difference between i and (k,j)

- Case 1: Moody's will inflate Goldman Sachs firms by 1 notch but not Berkshire Hathaway firms
- Case 2: Moody's will inflate Goldman Sachs firms by 1 notch and Berkshire Hathaway firms by 2 notch

	GS	BRK	Z	Rating 1	Rating 2
Issuer i	Y	Y	1	AA2	Aaa
Issuer j	Y	N	0	AA2	AA2
Issuer k	N	N	0	AA3	AA3
True β_0				0	2
Estimate $\hat{\beta}$				0.5	2.5

- **$\hat{\beta}$ will erroneously pick up Moody's bias towards other shareholders!!**

Econometric Motivation

Given a proxy Z for the cross-ownership between issuer i and the CRA, run: $Y_i = f(Z_i, control, \epsilon_i) = \beta_0 Z_i + \pi_0 controls + \epsilon_i$

Concern 1 **Z_i is difficult to find...**

- Kedia et al. (2016) employ a dummy variable approach: whether related to Warren Buffett or not
- Omitting other shareholders could bias the result

Concern 2 **What is the true rating model, $f(\cdot)$, to use?**

- misspecification bias

Concern 3 **β_0 has a average marginal effect (AME) interpretation**

- What about the heterogeneous impact?

Econometric Motivation

Given a proxy Z for the cross-ownership between issuer i and the CRA, run: $Y_i = f(Z_i, control, \epsilon_i) = \beta_0 Z_i + \pi_0 controls + \epsilon_i$

Concern 1 **Z_i is difficult to find...**

- Kedia et al. (2016) employ a dummy variable approach: whether related to Warren Buffett or not
- Omitting other shareholders could bias the result

Concern 2 **What is the true rating model, $f(\cdot)$, to use?**

- misspecification bias

Concern 3 **β_0 has a average marginal effect (AME) interpretation**

- What about the heterogeneous impact?

This paper

- I measure and track a bond issuer's relationship with over **1500** Moody's institutional shareholders in 2001 - 2016.
- Estimate a flexible, semi-structural bond rating model with information transmission
  - Propose a measure for heterogeneous marginal effect
  - Develop methodology for statistical inference

Model

The Econometric Model

Consider the following ordered-response model:

$$\text{Credit Ratings: } Y_i = \begin{cases} Aaa & \text{if } -\infty < y_i^* \leq c_1 \\ Aa & \text{if } c_1 < y_i^* \leq c_2 \\ \ldots & \\ C & \text{if } c_{L-1} < y_i^* \leq \infty \end{cases} \quad (1)$$

$$\textit{Latent Default Risk}: \quad y_i^* = f(\mathbf{X_i}, \mathbf{Z_i}, U_i) \quad (2)$$

Bahavioral Framework

- $\mathbf{c}_i$: a $(L-1)$-vector of unknown cutoff points between categories.
- $\mathbf{X_i}$: a vector that represents public information.
- $\mathbf{Z_i}$: a vector of shared-ownership relation measures
- U_i: an error term that represents private information.

Rating Probabilities

The probability that a bond will be rated better or equal to category K:

$$\begin{aligned} Prob(Y_i \leq K | X_i, Z_i) &\equiv Prob(y^* < c_K | X_i, Z_i) \\ &\equiv G_K(X_i, Z_i) \end{aligned} \quad (3)$$

But...

- Need large N to learn G_K
- Computation cost is heavy

Index Assumption

$$
\begin{aligned}
Prob(Y_i \leq K | X_i, Z_i) &\equiv Prob(y^* < c_K | X_i, Z_i) \qquad (4)\\
&\equiv G_K(X_i, Z_i) \\
&= H_K(\underbrace{X_{F1} + \mathbf{X_{F'}}\theta_0^F}_{\text{Firm Index}}, \underbrace{X_{B1} + \mathbf{X_{B'}}\theta_0^B}_{\text{Bond Index}}, \underbrace{Z_1 + \mathbf{Z}'\theta_0^Z}_{\text{CI index}})
\end{aligned}
$$

(K can be any rating category of interest, say AA)

Index Assumption

$$\begin{aligned} Prob(Y_i \leq K | X_i, Z_i) &\equiv Prob(y^* < c_K | X_i, Z_i) \qquad (4) \\ &\equiv G_K(X_i, Z_i) \\ &= H_K(\underbrace{X_{F1} + \mathbf{X_{F'}}\theta_0^F}_{\text{Firm Index}}, \underbrace{X_{B1} + \mathbf{X_{B'}}\theta_0^B}_{\text{Bond Index}}, \underbrace{Z_1 + \mathbf{Z}'\theta_0^Z}_{\text{CI index}}) \end{aligned}$$

(K can be any rating category of interest, say AA)

- Need to estimate both the **index parameters** $\theta_0 \equiv [\theta_0^F, \theta_0^B, \theta_0^Z]$ and the **link function** $H_K(\cdot, \cdot, \cdot)$

More on Identification | Comparison with Parametric and Nonparametric approaches

Marginal effects

- Once the index parameters are identified, so does the marginal effect of any x:

$$\Delta Pr(Y_i = j|x_1) = Pr(Y_i = j|x_1 + \Delta x) - Pr(Y_i = j|x_1)$$

Marginal effects

- Once the index parameters are identified, so does the marginal effect of any x:

$$\Delta Pr(Y_i = j|x_1) = Pr(Y_i = j|x_1 + \Delta x) - Pr(Y_i = j|x_1)$$

- I consider estimation and inference of a local objects termed

Quantile Marginal effects (QME^j_q): $\quad = E[\Delta Pr(Y_i = j|x_1)|x_2 \in \mathbb{Q}_{x_2}]$

  - x_1 and x_2 can be the same
  - $x_2 \in \mathbb{Q}_{x_2}$: a specific x-quantile of interest
  - Example: whether the impact of financial leverage is different for small, mid and large-cap companies

Estimation and Inference

Two-stage Estimation

Stage I Semiparametric ML estimator for θ_0

- Likelihood function: $\Pi_i P_{1i}^{Y_{1i}}(\theta) P_{2i}^{Y_{2i}}(\theta) \cdots P_{Li}^{Y_{Li}}(\theta)$
- Log-Likelihood function:

$$\hat{Q}(\theta) = \frac{1}{N} \sum_i \{ \sum_{k=1}^{L} Y_i^k Log(\widehat{P_i^k}(\theta)) \} \quad (5)$$

- A kernel estimator for $P_i^k(\theta)$: the probability that bond i will be rated as category k

Stage II Estimate the quantile marginal effects

A kernel Estimator for $P_i^k(\theta)$

- Let V_i be the index vector at a sample point i and v be a point of interest, I use Shen and Klein (2017)'s estimator:

$$\widehat{P^*_{t+1}(v,\theta)} := \frac{\sum_i (Y_i^k - \widehat{\Delta^t_{iv}}) K_h(V_i - v)}{\sum_i K_h(V_i - v)} \tag{6}$$

  - The kernel function $K_h(\cdot) \equiv \frac{1}{Nh}\phi(\frac{V_i - v}{h})$
  - Intuition: when estimating P_v, K_h assigns lower weight for observations V_i that are "different" from v.
  - $\widehat{\Delta^t_{iv}} \equiv \widehat{P^*_{t-1}(V_i,\theta)} - \widehat{P^*_{t-1}(v,\theta)}$ reduces the bias
  - apply iteratively $t = 1, 2, ...$ to ensure normality when $dim(v)$ is large
- I formally prove the asymptotic distribution of estimated parameters and marginal effects given this proposed estimator

Detail

Asymptotic Results I

Theorem (Normality of index parameters)

Base the estimation on $\widehat{P^*(v,\theta)}$*, the bias corrected estimator of conditional expectation, with appropriate bandwidth h,*

$$\sqrt{N}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, H_0^{-1} \Sigma H_0^{-1})$$

Stage II: Estimating marginal effects

The "true" quantile marginal effect: $QME_q^j \equiv E[\Delta Pr(Y_i = j|x)|x_2 \in \mathbb{Q}_{x_2}]$
Its sample analogue:

$$\widehat{QME_q^j} = \frac{\sum_{i=1}^{N} \hat{t}_{qi} \Delta Pr\widehat{(Y_i = j}|x_1)}{\sum_{i=1}^{N} \hat{t}_{qi}} \tag{7}$$

1. Estimate $\Delta Pr\widehat{(Y_i =} j|x_1) = Pr(Y_i \widehat{= j|x_1} + \Delta x) - Pr\widehat{(Y_i = j}|x_1)$
2. Estimate Quantile: $\hat{t}_{qi} \equiv 1\{F_N^{-1}(q_1) < x_{2i} < F_N^{-1}(q_2)\}$
   - F_N : empirical CDF of x_2
   - (q_1, q_2) defines the quantile of interest

Asymptotic Results II

Theorem

Under A.1-A.5 in the paper, with the quantile marginal effect QME_q^j and its estimator defined above, we have

$$\sqrt{N}(QME_q^j - \widehat{QME_q^j}) \sim N(0, \sum_{k=1}^{j-1} E[\psi_j' \psi_j])$$

where $\psi_j^k \equiv \psi_{1i}^k + \psi_{2i}^k + \psi_{3i}^k + \psi_{4i}^k$

- ψ_{1i}^k: parameter estimation uncertainty of θ
- ψ_{2i}^k: quantile estimation uncertainty of $\hat{t}_{qi}$
- ψ_{3i}^k: uncertainty due to the increment Δx

Detail on the Asym Covariance Matrix

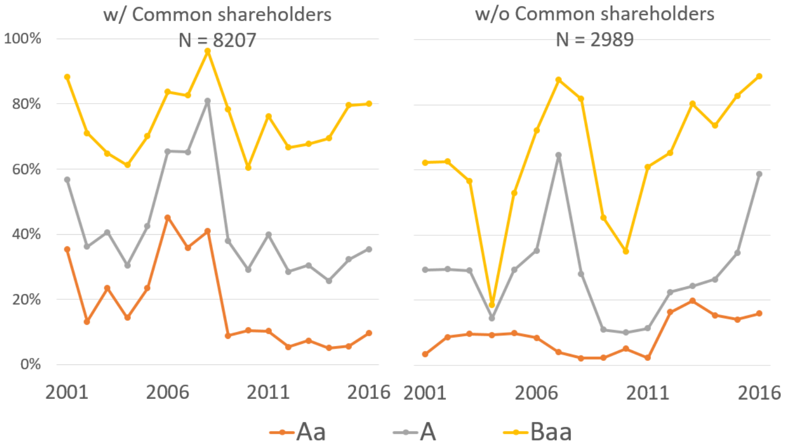
Empirical Results

Institutions & Data

- Rating data: Mergent's Fixed Income Securities Database (FISD). 11364 new bonds by 1462 firms
- Firm + bond characteristics (X_i): CRSP-Compustat quarterly file and FISD
- Sampling starts from 2001 to 2016
- Firm-Moody relations (Z_i): Thomson-Reuters Institutional Holdings (13F) Database

Summary Statistics: Ratings

Figure 1: Moody's Ratings



y-axis: proportion of ratings that are in category X or better

Moody's large shareholders and their ownership stakes

Shareholder	T	Mean	Max	Min
HARRIS ASSOCIATES L.P.	21	2.42%	5.02%	0.00%
CHILDREN'S INV MGMT (UK) LLP	20	2.29%	5.31%	0.01%
SANDS CAPITAL MANAGEMENT, INC.	28	3.01%	5.59%	0.40%
T. ROWE PRICE ASSOCIATES, INC.	64	1.47%	5.94%	0.18%
BARCLAYS BANK PLC	55	2.52%	6.32%	0.03%
GOLDMAN SACHS & COMPANY	63	1.94%	7.24%	0.01%
VALUEACT CAPITAL MGMT, L.P.	13	5.19%	7.77%	0.93%
VANGUARD GROUP, INC.	64	3.79%	7.98%	1.64%
MSDW & COMPANY	57	2.20%	8.14%	0.22%
DAVIS SELECTED ADVISERS, L.P.	51	5.56%	8.14%	0.10%
FIDELITY MANAGEMENT & RESEARCH	64	1.99%	9.08%	0.00%
CAPITAL RESEARCH GBL INVESTORS	13	4.80%	11.31%	0.07%
CAPITAL WORLD INVESTORS	35	6.07%	12.60%	0.66%
BERKSHIRE HATHAWAY INC.	64	14.87%	20.43%	11.33%

T: Periods of holding Moody's stocks (out of 64 quarters)

Characterizing Conflicts of Interest

Z_1: Overlapping shares, defined as the total percentage of Moody's stock owned by all common shareholders (*o-share*).

Z_2: Number of Large Shareholder[1] (*num_largeSH*).

Z_3: Number of common shareholder (*num_SH*).

I define "Moody-Firm-Ownership-Interaction"(MFOI) to be

$$MFOI_i = Z_1 + \theta_1 Z_2 + \theta_2 Z_3 \tag{8}$$

- no need to specify θ_1, θ_2

[1] We follow Kedia et al. (2016) and pick 5% as the threshold point

Summary Statistics

Table 1: Covariates

Variable	Description	Mean	Std. Dev.	Min	Max
ASSET	log(asset) of the issuer	9.958	2.134	4.360	14.935
STABILITY	variance of asset	0.174	0.156	0.000	1.504
LEVERAGE	Leverage ratio	0.258	0.165	0.000	1.283
PROFIT	profit/sales	0.035	0.058	-0.681	0.503
AMT	log(issuing amount)	12.695	1.457	2.708	19.337
SENIORITY	subordination status	0.859	0.347	0.000	1.000

Table 2: Shared-ownership Relation

Variable	Description	Mean	Std. Dev.	Min	Max
largeSH	number of large shareholders[2]	1.019	0.802	0	5
SH	number of shareholders	211.618	107.231	0	440
oshare	total ownership stakes	0.459	0.142	0	0.988

[2] more than 5% ownership stake in Moody's

Empirical Model

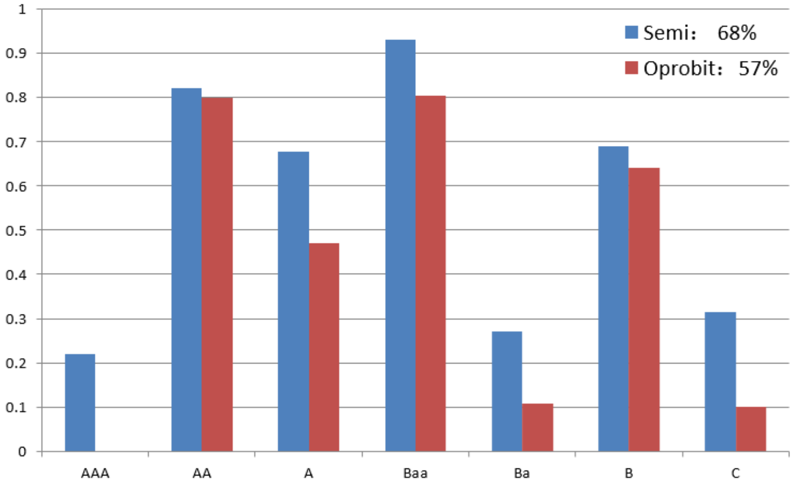
With the following index structure defined,

$$\begin{aligned} V_F &= ASSET + \theta_1^F STABILITY + \theta_2^F LEVERAGE + \theta_3^F PROFIT \\ V_B &= AMT + \theta_1^B SENIORITY \\ MFOI_i &= OSHARE + \theta_1 largeSH + \theta_2 SH \end{aligned}$$

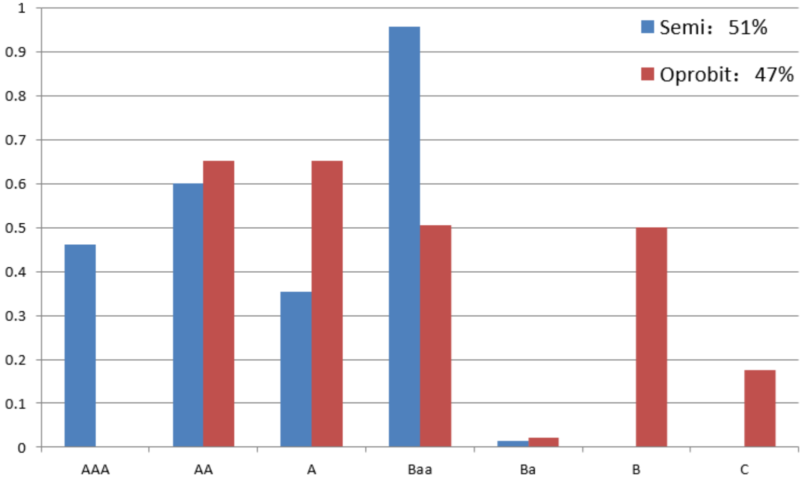
I use Moody-Firm-Ownership-Index (MFOI) to capture the cross-ownership and estimate

- **semiparametric model:** $Pr(Y \leq j|X, R) = P_j(V_F, V_B, MFOI)$
- **parametric ordered-probit model:** $Pr(Y \leq j|X, R) = \mathbf{\Phi}(T_j - \theta_0^F V_F - \theta_0^B V_B - \theta_0 MFOI)$

Prediction Results: 2001-2007 (in-sample)



Prediction Results: 2009-2015 (out of sample)



Index Parameters and Average Marginal effects

		Marginal Effects (percentage point)						
	θ	AA	A	Baa	Ba	B	C	Average
Semiparametric								
Asset***	1.00	0.11	1.71	5.89	6.13	4.31	0.64	4.42
stability***	-2.71	-0.23	-4.46	-8.89	-10.80	-8.81	-1.25	-5.74
leverage***	-4.25	-0.01	-0.97	-2.08	-2.78	-2.57	-0.48	-1.49
profit***	24.21	0.49	4.81	17.88	15.52	10.06	1.31	9.91
AMT***	0.41	0.05	0.07	-0.09	-1.96	-1.23	0.08	-0.49
seniority***	1.00	0.81	0.62	3.36	8.52	4.52	-0.22	3.10
MFOI***	1.00	0.51	9.78	9.00	2.12	2.27	0.14	5.86
Ordered-probit								
(with Year and Industry Fixed Effects)								
Asset***	1.00	0.98	5.69	9.12	10.05	7.47	1.69	6.77
stability***	-0.51	-0.67	-4.35	-5.91	-5.83	-4.45	-1.07	-3.71
leverage***	-5.14	-0.41	-2.65	-3.59	-3.54	-2.71	-0.65	-2.26
profit***	14.92	1.83	11.87	16.11	15.90	12.14	2.91	10.13
AMT	-0.09	-0.02	-0.11	-0.15	-0.18	-0.13	-0.03	-0.06
seniority***	1.00	1.08	6.45	8.75	8.63	6.59	1.58	5.48
MFOI***	-71.13	0.51	3.02	4.84	5.33	3.96	0.89	3.59

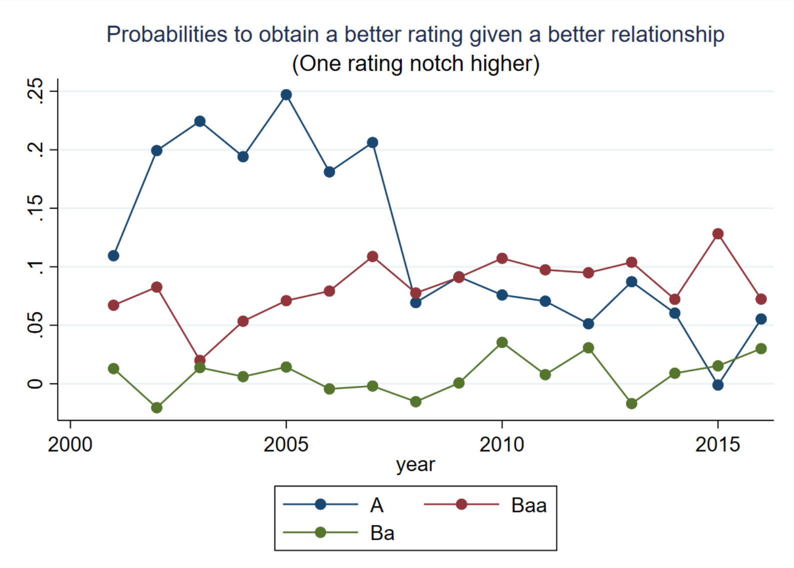
Rating Transition Probabilities

	Original rating category				
	Aa	A	Baa	Ba	B
Aaa	2.65%	2.64%	0.74%	0.18%	0.01%
Aa	**15.87%**	**16.64%**	5.58%	2.04%	0.53%
A	**-11.40%**	-3.05%	**9.32%**	4.79%	2.26%
Baa	-6.38%	**-12.85%**	**-7.41%**	1.75%	2.79%
Ba	-0.56%	-2.39%	-5.23%	-4.66%	-2.15%
B	-0.15%	-0.91%	-2.73%	-3.70%	-3.07%
C	-0.02%	-0.08%	-0.27%	-0.41%	-0.38%

- I report the pairwise transitional probability from a one standard deviation change in *MFOI*
- Numbers in **bold** are significant at 95% level.
- *Baa* and *A* bonds are affected by the most.

Time Variation

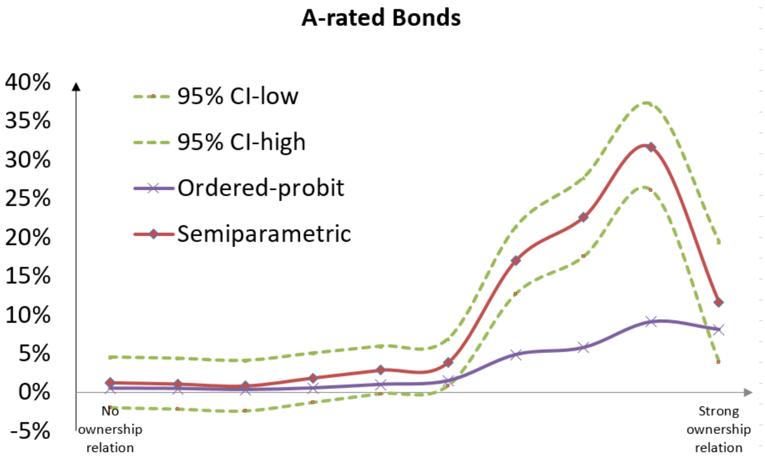
- Run the same model in each cross-section and compute the average marginal effect



Hypotheses

1. CRAs may assign favorable ratings to issuers related with their large shareholders. (Kedia et al., 2016)
2. For issuers with same level of conflicts of interest, the CRA is more conservative on low quality bonds. (An implication of Bolton et al. (2012))

Investment-grade bonds



- Moody's favorable treatment becomes statistically significant only for firms that are well connected with its shareholders
- and not monotonic...

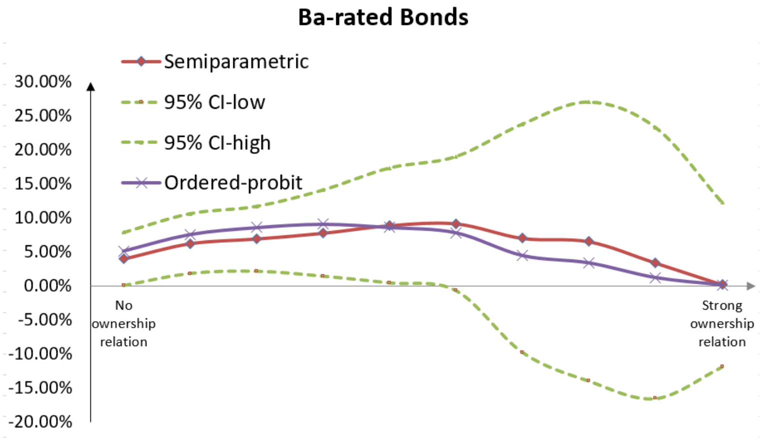
The big dip in the last decile: Warren Buffett effect?

Table 3: percentage of bonds related with major shareholders

Shareholders	Q7	Q8	Q9	Q10
VANGUARD GROUP, INC.	94.92%	98.04%	89.66%	98.66%
DAVIS SELECTED ADVISERS	15.78%	26.65%	38.77%	56.86%
FIDELITY	96.70%	98.57%	91.35%	97.50%
CAPITAL RESEARCH	11.59%	13.01%	10.61%	13.99%
BERKSHIRE HATHAWAY INC.	2.05%	5.44%	7.49%	39.04%

- Unlike other major shareholders, Berkshire Hathaway disproportionately owns more firms in Q10
- Anecdotal evidence: Moody's was slow in downgrading Wells Fargo and Bank of America, two other firms that Buffett owns Kedia et al. (2016)
- Public concern on Warren Buffett large holding on Moody's

High-Yield bonds



- Moody's favorable treatment is less significant on low credit quality bonds
- Overrating a high yield bond puts Moody's reputation in greater danger...


YOU CAN PUT LIPSTICK ON A PIG, BUT IT'S STILL A PIG

Conclusions and Future Research

- Propose a semiparametric multi-index model to study the rating process
- Shared-ownership has heterogeneous impact on ratings
  - Effects vary from 0 (statistical sense) to 31%***
  - Issuers with a stronger connection with Moody's (a larger $MFOI_i$) are more likely to receive higher ratings.
  - Low quality bonds, regardless the CRA-issuer relation, do not receive much favorable treatment.
- **Ongoing and future research**
  - High-dimensional/variable selection method
  - Cross-market impact of credit ratings: Jiang and Mizrach (2018)

Thank you very much!

Questions?

yixiao@economics.rutgers.edu

Prediction accuracy

	In sample (2001-2007)			Out of sample (2009-2016)		
	Semipara.	Oprobit	N	Semipara.	Oprobit	N
AAA	22.03%	0.00%	59	46.03%	0.00%	63
AA	82.21%	79.91%	916	60.14%	65.19%	296
A	67.70%	47.10%	1446	35.33%	65.15%	1101
Baa	93.00%	80.41%	1442	95.75%	50.54%	2025
Ba	27.17%	10.84%	784	1.44%	2.15%	626
B	69.00%	63.99%	958	0.00%	50.00%	762
C	31.43%	10.00%	140	0.00%	17.65%	158
Total	69.70%	57.17%	5745	50.57%	46.83%	5031

Stylized Behavioral Framework

- A bond's default risk y^* is driven by a vector of observed bond+firm characteristics x and unobserved component U: $y^* = x^T\beta_0 + U$
- Bond issuers and the CRA have **common shareholders**.
- The institutional cross-ownership, or the "issuer-CRA relationship", is captured by $\mathbf{Z_i}$ of arbitrary dimension
  - $\mathbf{Z_i}$ is common knowledge
- I show that, in equilibrium, CRA's estimate of default risk y^* is a **non-separable** function of Z and U:

$$y^* = x^T\beta_0 + H(Z, U)$$

go back

Literature

- **Credit rating model**: Blume et al. (1998); Horrigan (1966); Kaplan and Urwitz (1979); West (1970)
- **Rating bias**: Kedia et al. (2017); Becker and Milbourn, (2011); J. X. Jiang et al., (2012)
- **Endogeneity in semiparametric model:** Blundell and Powell (2004); Imbens and Newey (2009); Chesher and Smolinski (2012); Chesher, Rosen and Smolinski (2013)
- **Threshold estimation**: Manski (1985); Horowitz (1992); Lewbel, 1997, 2000; R. W. Klein and Sherman (2002)

go back

Identification Assumptions

1. **Assumption** $\mathcal{I}1$ $H_K(\cdot,\cdot,\cdot)$ is differentiable and monotonic in each index on a set with positive probability.
2. **Assumption** $\mathcal{I}2$ All three indices have at least one continuous covariate that is not in the other two.
3. **Assumption** $\mathcal{I}3$ **(uniqueness)** For any $(F_i, B_i, Z_i) \in \mathcal{A}$, a set with probability, $Pr(Y_i \leq K|V_F(\theta^*), V_F(\theta^*), Z_i) = Pr(Y_i \leq K|V_F(\theta_0), V_F(\theta_0), Z_i) \implies \theta^* = \theta_0$

go back

Other approaches

- Nonparametric approach:
  - Directly estimate $Prob(Y_i \leq K|W_i) = F_K(X_i, Z_i)$
  - perform well only in really large samples ("curse of dimensionality")
- Parametric approach:
  - Assumes $Prob(Y_i \leq K|W_i) = \Phi(c_K - W_i\beta_0)$
  - w/ Φ the cdf of standard normal r.v
  - Perform well in small sample, but
  - not valid if the latent default risk $y^* \neq X_i\beta_0 + U_i$ or U_i is not $N(0, \sigma^2\mathcal{I})$
- The semiparametric approach takes a middle ground, balancing **estimation efficiency** and **robustness**

go back

Bias Correction

- In a large class of semiparametric index models, $\sqrt{N}(\hat{\theta} - \theta_0)$ depends on the score (or gradient):

$$\sqrt{N}\hat{G}(\theta_0) = \mathcal{A}(P, \hat{w}) + \mathcal{B}ias(\hat{P}^*, \hat{w})$$

where the estimated weight function $\hat{w}_i = \alpha_i \nabla_\theta \widehat{P}(v_i, \theta_0)$ and

$$\begin{aligned}
\mathcal{A}(P, \hat{w}) &\equiv N^{-1/2} \sum_i \tau_i (Y_i - P_i)\hat{w} = \overbrace{\mathcal{A}(P, w)}^{apply\ CLT} + o_p(1) \\
\mathcal{B}ias(\hat{P}^*, \hat{w}) &\equiv N^{-1/2} \sum_i \tau_i (\widehat{P_i^*} - P_i)\hat{w} \\
&= \underbrace{N^{-1/2} \sum_i \tau_i (\widehat{P_i^*} - P_i) w \frac{\hat{g}}{g}}_{\mathcal{B}ias(\hat{P}^*, w)} + o_p(1)
\end{aligned}$$

g here is the joint density of the index vector v.

- The recursive structure of $\widehat{P^*(v, \theta)}$ makes $\mathcal{B}ias(\hat{P}^*, w)$ difficult to analyze.
- Strategy: replace it with another object that is easier to study

A U-statistics equivalence result

From before,

$$\sqrt{N}\hat{G}(\theta_0) = \mathcal{A}_{iid} + \mathcal{B}ias(\hat{P}^*, w) + o_p(1)$$

Then I show

$$\mathcal{B}ias(\hat{P}^*, w) - \mathcal{B}ias(\hat{P}, w) = o_p(1)$$

where $\hat{P} \equiv \frac{\sum_i Y_i^K K_h(V_i - v)}{\sum_i K_h(V_i - v)}$ is the regular Nadaraya-Watson kernel estimator for conditional expectations.

- Importantly, I use a "residual property" of semiparametric derivatives:

$$E[w|v_i(\theta)]_{\theta=\theta_0} = 0$$

- $\mathcal{B}ias(\hat{P}, w)$ is a degenerate $U-$statistic. Therefore easy to show $o_p(1)$ (Sherman, 1994; Nolan and Pollard, 1987)

go back