Selecting Directors Using Machine Learning

Isil Erel, Léa H. Stern, Chenhao Tan, Michael S. Weisbach
Working Paper

王健
2020-05-13

content

- Introduction
  - Background & Motivation
  - Question
  - Research content
  - Related researches
  - Contribution
- Data & Variables
- Empirical results
- Conclusion

1. Introduction

Background & Motivation

- In practice, there is much variation in director quality and the extent to which they serve shareholders' interests.
- The selection process for selecting directors is one of the most important yet least studied questions in corporate governance.
- While "traditional" econometrics is typically designed for estimating structural parameters and drawing causal inferences, machine learning is substantially better at making predictions, in part because it does not impose unnecessary structure on the data.

1. Introduction

Question: Can machine learning algorithms help address director selection problem in corporate governance?

- We outline a potential way in which they can.
- We argue that algorithms can shed light on the decision-making process that governs the nomination of corporate directors.

1. Introduction

Research contents

- We design a process for assessing a director's performance in a particular firm.(shareholder votes)
- We construct machine learning algorithms to predict the performance of any potential director at any particular company, taking into consideration who is currently sitting on the board.("selective labels" problem, recommendation problem)
- We use the algorithm's predictions to understand the features that are overvalued and undervalued by firms in the director selection process.

1. Introduction

Constructing a sample on which algorithms can select directors.

↓

Evaluating machine learning predictions of director performance.

- *Predictions of Director Performan*
- *Excluding Poorly Performing Firm*

↓

Designing the quasi-Labels procedure to evaluate the algorithm.

"selective labels" problem

↓

Director popularity or performance?

recommendation problem

↓

Characteristics that affect director performance.

- *Univariate Comparisons*
- *Multivariate Comparisons*
- *Overvalued Director Characteristi*

1.Introduction

Related researches

- Cai et al. (2009), Fischer et al. (2009), and Iliev et al. (2015) suggest that cross-sectional variation in shareholder support does in fact reflect market perceptions of director quality.
- We develop a quasi-labels procedure framework in the spirit of Kleinberg et al. (2017)
- Ertimur et al. (2018) report that since 2003 large institutional investors take an active role in developing the guidelines that are the basis of ISS recommendations.
- Aggarwal et al. (2016) find that shareholders are less likely to follow the recommendations.
- Iliev Lowry (2014) show that institutional investors with larger size of ownership tend to vote more independently from ISS recommendations.

1.Introduction

Contribution

- We present a new approach based on algorithms to select the directors of publicly traded companies and our finding on the predictability of which directors will or will not be popular with shareholders has important implications for corporate governance.
- Paper prove that the shareholders can judge the directors independently, and excess votes can be a good variable to evaluate the performance of individual director.
- Paper confirm that the board selection process leads to directors who often those nearest at hand and are not necessarily the best choices to serve shareholders' interests, and, provide us with the tools to change it.

2. Data & Variables

1) Data: Sample Selection

Director and board-level characteristics

- Sources: BoardEx
- Time :2000-2014
- Sample size: 41015(directors) and 4887(firms)

Shareholder support for individual directors

- Sources: ISS Voting Analytic

Firm and industry-level features

- Sources: Compustat /CRSP

Announcement dates

- Sources: BoardEx, CapitalIQ and Lexis-Nexis

Data description

Director-level features: 64

A.4.1. Individual Director Features

Source: BoardEx except if stated otherwise
(as of when the director joins the board)

Variable	Definition
Age	Director age
Audit chair	Equals one if director is chair of the audit committee
Audit member	Equals one if director is a member of the audit committee
Avgtimeothco	The average time that a director sits on the board of quoted companies
Bkgd academic	Equals one if job history includes in title one of the following: "professor" "academic" "lecturer" "teacher" "instructor" "faculty" "fellow" "dean" "teaching"
	Equals one if job history includes in title one of the following: "underwriter" "investment" "broker" "banker" "banking" "economist"

Board-level features: 56

A.4.2. Board-level features

Source: BoardEx except if stated otherwise
(as of when the director joins the board)

Variable	Definition
Attrition rate	Number of Directors that have left a role as a Fraction of average number of Directors for the preceding reporting period
Average age	Average age of directors on the board

Firm-level features: 76

Industry-level features: 8

A.4.3 Firm level features

Source: Compustat /CRSP except if stated otherwise
(as of when the director joins the board)

Variable	Definition
Current assets	Current assets - Total
Acquisitions	Acquisitions
Auditor	Dichotomous variable for each auditing firm
BCW	Equals one if firm was on the Fortune-Best Company to work for list within 10 years preceding the nomination (from Alex Edmans' website)

2. Data & Variables

2) Variables: Re-Election Results

- Excess votes:

$$excess\ vote_i = \frac{1}{3}\sum_{t=1}^{3}(support\ rate_{i,t} - \overline{support\ rate_t})$$

Two concerns:

- In the vast majority of cases, directors receive overwhelming majority.
- Votes could reflect arbitrary recommendations by proxy advisors such as ISS.

2. Data & Variables

2) Summary Statistics

	n	mean total votes	median total votes	mean excess votes	std excess votes	25th ptcl excess votes	median excess votes	75th pctl excess votes
2000	331	0.950	0.974	0.0008	0.0300	-0.0058	0.0004	0.0082
2001	772	0.944	0.970	-0.0001	0.0455	-0.0050	0.0017	0.0134
2002	1,057	0.946	0.970	0.0022	0.0387	-0.0038	0.0015	0.0115
2003	1,774	0.951	0.974	0.0064	0.0359	-0.0014	0.0028	0.0149
2004	2,019	0.953	0.977	0.0069	0.0442	-0.0008	0.0033	0.0153
2005	1,893	0.948	0.974	0.0049	0.0369	-0.0011	0.0033	0.0136
2006	1,789	0.941	0.969	0.0051	0.0412	-0.0016	0.0036	0.0153
2007	1,942	0.940	0.971	0.0045	0.0434	-0.0023	0.0026	0.0157
2008	1,691	0.944	0.973	0.0067	0.0431	-0.0032	0.0034	0.0180
2009	1,541	0.948	0.976	0.0072	0.0435	-0.0020	0.0045	0.0187
2010	1,842	0.948	0.977	0.0039	0.0431	-0.0044	0.0027	0.0152
2011	1,825	0.954	0.981	0.0038	0.0462	-0.0019	0.0035	0.0160
2012	1,862	0.952	0.981	0.0045	0.0422	-0.0007	0.0038	0.0162
2013	2,148	0.948	0.980	0.0027	0.0444	-0.0021	0.0032	0.0139
2014	1,568	0.959	0.985	0.0063	0.0408	-0.0004	0.0045	0.0149
	24,054	0.9484	0.9755	0.0044	0.0413	-0.0024	0.0030	0.0147

2. Data & Variables

2) Summary Statistics

Table illustrates that the frequency of shareholder discontent varies by director and board characteristics

However, theory provides little guidance of the relation between the various director, board and firm characteristics and the performance of directors.

	Full sample	yes	no	Difference p-value
Director level				
Male	0.102	0.106	0.079	0.000
Foreign	0.101	0.115	0.100	0.138
Qualifications > median	0.102	0.094	0.106	0.005
Network size > median	0.102	0.108	0.096	0.002
Generation BBB	0.101	0.093	0.118	0.000
Generation X	0.101	0.151	0.096	0.000
Busy director	0.102	0.145	0.090	0.000
Finance background	0.102	0.106	0.101	0.328
Board level				
Fraction male > median	0.102	0.116	0.091	0.000
Board size > median	0.102	0.089	0.114	0.000
Nationality mix > median	0.102	0.108	0.100	0.064
Attrition rate > median	0.098	0.106	0.086	0.000

3. Empirical results

1) Machine Learning Predictions

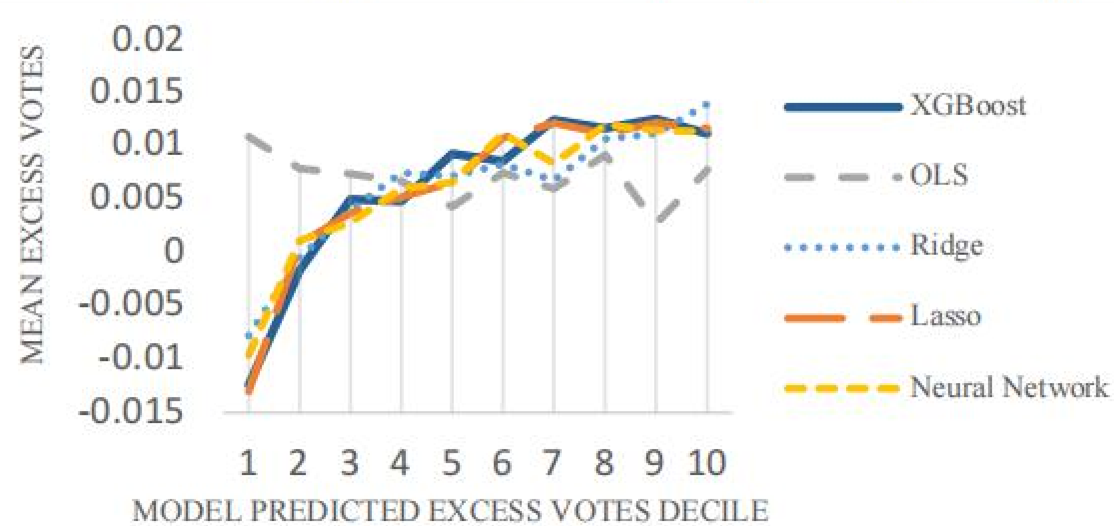
- **Train set:** each algorithm on the 2000-2011 portion of our sample containing 18,476 new independent director appointments, of which 12,815 are unique directors, at 2,407 firms.
- **Test set:** 2012-2014 portion of our sample containing 5,578 new director appointments, of which 4,019 are unique directors, at 569 firms.
- **Models:** OLS, XGBoost, Ridge, Lasso, Neural Network

3. Empirical results

The figure shows how the mean shareholder support for a director is an increasing function of the predicted one for all the machine learning algorithms, but not for the OLS model.

Average Observed Performance for Directors in a Given Percentile of Predicted Performance as Predicted by:

	Predicted Percentile of Excess Votes	OLS	XGBoost	Ridge	Lasso	Neural Network
Directors predicted to perform **poorly**	1%	0.028	-0.031	-0.012	-0.024	-0.014
	5%	-0.018	-0.014	-0.013	-0.015	-0.010
	10%	0.014	-0.008	0.000	-0.008	-0.001
Directors predicted to perform **well**	90%	0.013	0.013	0.011	0.011	0.011
	95%	0.007	0.012	0.014	0.013	0.016
	100%	0.006	0.011	0.009	0.016	0.015



3. Empirical results

2) Excluding Poorly Performing Firms

- A possible concern with this analysis is that the relation between predicted performance and subsequent performance could occur only because of poorly performing firms.
- A poorly performing firm would likely be less attractive to a director, so it could be that only low ability directors are attracted to poorly performing firms, even if the firms are relatively large and otherwise prestigious.
- We repeat our analyses omitting firms that experience negative abnormal returns in the year prior to the nomination and find similar results without poorly performing firms in the sample.

3. Empirical results

3) Designing the Quasi-Labels Procedure

- Accurate out of sample predictions, however, are not sufficient to imply that algorithms could assist firms in their nominating decisions of corporate directors.

Two concerns:

- selective labels problem
- when deciding on their choice of directors, decision makers presumably take factors into account that are not observable to the algorithm.

Methodology:

- To formalize these concepts, we design a pool of realistic potential candidates for each vacant board position.

3. Empirical results

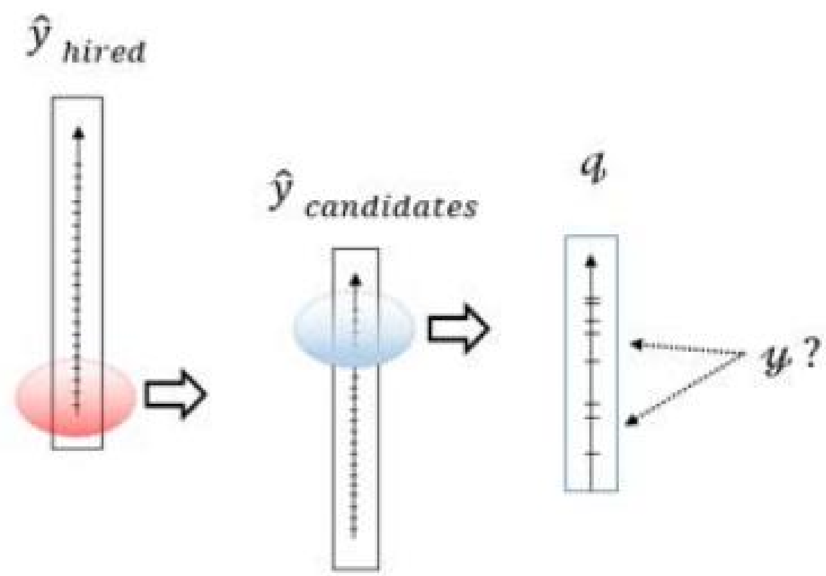
3) Designing the Quasi-Labels Procedure

Pool design details:

- Each new board appointment in the test set is associated with a candidate pool, comprised of directors who, within one year of the appointment, joined the board of **a smaller neighboring** firm.
- A neighboring firm is defined as a firm whose headquarters is within 100 miles of the focal firm’s headquarters. The average distance with the focal board is 35 miles (median distance is 26 miles).
- A smaller neighboring firm(147) & in the same industry(33)
- In the rest of the paper, we focus on results with XGBoost to simplify the discussion.

3. Empirical results

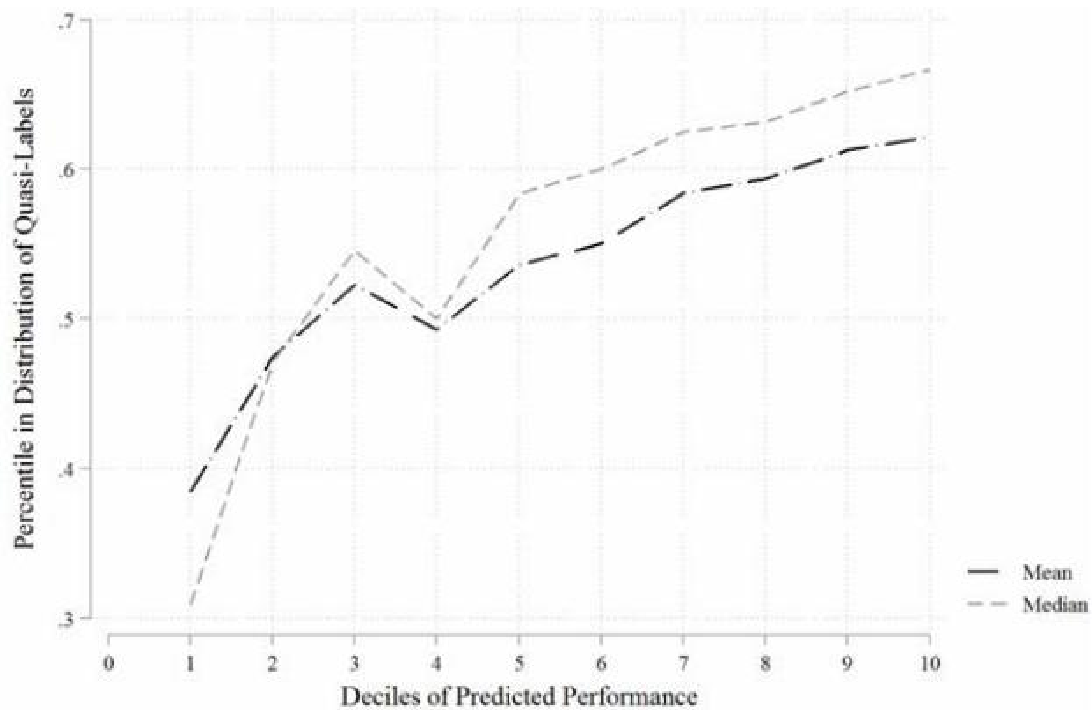
3) Designing the Quasi-Labels Procedure



3. Empirical results

	Median percentile of observed performance in the distribution of quasi-labels (candidate pools)				
	OLS	XGBoost	Ridge	Lasso	Neural Network
Bottom decile of predicted performance	77th	27th	37th	23rd	29th
Top decile of predicted performance	75th	78th	82nd	79th	69th

In our setting, the endogenous nature of the board-director match could lead to **systematically inflated** quasi-labels, i.e. by revealed preference, the performance of the available candidate would not be as high on the focal board.



3. Empirical results

4) Director Popularity or Performance?

- Many institutional shareholders decide on their votes through recommendations of shareholder services companies such as ISS.
- The recent literature on routine director re-election does report that votes indeed capture the performance of directors, we test whether we find evidence for this in our data.
- Specifically, we compare the **cumulative abnormal returns** (CARs) around the announcement of director appointments in our test set for directors predicted to do well to those for directors predicted to do poorly.
- Reports the mean CARs using a **(-1; +1) window** around announcements.

3. Empirical results

4) Director Popularity or Performance?

	N	Mean	Median
Directors in Decile 1 of predicted performance (excess votes)	292	-1.94%	-0.64%
Directors in Decile 10 of predicted performance (excess votes)	575	0.75%	0.34%
Difference in means (p-value)		**0.0043**	

Use our XGBoost algorithm to predict excess votes (and also used the algorithm to predict announcement CARs)

3. Empirical results

firm profitability three years after the director has been appointed (EBITDA/Total Assets)

4) Director Popularity or Performance?

		1	2	3	4	5	6	7	8	9	10	Decile 10 - 1 *p-value*
Algorithm trained on **profitability**	Average observed profitability	-0.498	-0.064	-0.017	0.017	0.078	0.083	0.113	0.114	0.144	0.205	0.0000
	Average observed shareholder support	0.942	0.946	0.956	0.937	0.957	0.961	0.953	0.954	0.960	0.961	0.0002
	Average observed excess votes	-0.0004	0.002	0.006	0.002	0.006	0.004	0.003	0.005	0.006	0.004	0.0668
Algorithm trained on **excess votes**	Average observed profitability	0.006	-0.017	0.008	0.037	0.052	0.058	0.057	0.083	0.087	0.112	0.0000
	Average observed excess votes	-0.012	-0.002	0.005	0.005	0.009	0.009	0.012	0.012	0.013	0.011	0.0000
Algorithm trained on **total votes**	Average observed profitability	-0.003	-0.032	-0.031	-0.018	0.024	0.029	0.058	0.075	0.086	0.100	0.0000
	Average observed shareholder support	0.920	0.937	0.946	0.948	0.950	0.957	0.957	0.966	0.972	0.977	0.0000

The model trained to predict profitability in the subsequent period indeed does predict future profitability well. What is perhaps more surprising is that even though the model is trained to predict profitability, it can also predict future shareholder support.

3. Empirical results

5) Characteristics that Affect Director Performance

Univariate Comparisons

	Mean		
	Bottom decile of predicted performance	**Top** decile of predicted performance	Difference p-value
Busy	0.520	0.120	0.000
Chairman	0.098	0.001	0.000
Compensation committee	0.624	0.059	0.000
Compensation committee chair	0.175	0.024	0.000
Foreign	0.156	0.088	0.005
Governance chair	0.045	0.011	0.000
Governance committee	0.168	0.122	0.008
International work experience	0.109	0.037	0.000
Male	0.897	0.746	0.000
Network size	1540	1327	0.000
Nomination chair	0.004	0.001	0.318
Nomination committee	0.023	0.011	0.057
Number of qualifications	2.208	2.282	0.180
Total current number of boards sitting on	2.848	1.545	0.000
ROE	-0.110	0.194	0.353
Stock returns prior 12 months	0.158	0.116	0.188

Top decile means good predicted performance

3. Empirical results

5) Characteristics that Affect Director Performance

Multivariate Comparisons

Dependent variable: predicted performance	(1)	(2)	(3)	(4)
Busy	-0.006***	-0.005***	-0.005***	-0.005***
	(-24.332)	(-13.183)	(-12.230)	(-12.087)
Male	-0.001***	-0.001	-0.001	-0.001*
	(-4.623)	(-1.398)	(-1.603)	(-1.688)
Age		-0.000**	-0.000**	-0.000**
		(-2.001)	(-2.079)	(-2.242)
MBA		0.000	0.000	0.000
		(1.074)	(1.108)	(1.150)
Ivy league		-0.001**	-0.001*	-0.001*
		(-2.555)	(-1.869)	(-1.864)
Background lawyer			-0.002	-0.001
			(-1.521)	(-1.394)
Background academic			0.000	0.000
			(0.180)	(0.134)
Background finance			-0.001	-0.001
			(-1.344)	(-1.568)
Network size			-0.000***	-0.000***
			(-2.838)	(-2.691)
Ln (Assets)	0.001***	0.000	0.000	0.000
	(9.054)	(0.276)	(0.016)	(0.160)
ROA	0.001***	0.000	0.000	0.000
	(4.280)	(0.024)	(0.156)	(0.124)
Board size	-0.000***	-0.000**	-0.000*	-0.000*

3. Empirical results

5) Characteristics that Affect Director Performance

Overvalued Director Characteristics

- The predictions can help us identify the individual director features that tend to be overvalued or undervalued by firms when they select new directors.
- To do so, we identify directors who were nominated but were of predictably low quality and we compare them to those directors the algorithm would have preferred for that specific board position.
- These results highlight the features that are likely overrated by management when nominating directors. They are consistent with the view that directors tend to come from an “old boys club”, in which men who have sat on a lot of boards are chosen to be directors.

3. Empirical results

Compared to promising candidates identified by the algorithm, predictably unpopular directors are on average more likely to be **male**, have fewer degrees post undergraduate, a larger professional network, more current and past directorships, and are more likely to have a background in finance

	Hired directors with predicted and observed low shareholder support	Promising candidates for this board position	
	Mean	Mean	Difference *p-value*
Male	0.984	0.835	0.000
Number of qualifications	2.1	2.4	0.000
Ivy League	0.29	0.26	0.523
MBA	0.57	0.38	0.000
Network size	1673	1428	0.000
Total number of listed boards sat on	6.4	2.3	0.000
Total number of unlisted boards sat on	11.0	2.7	0.000
Total current number of boards sitting on	3.1	1.5	0.000
Number previous jobs same industry	0.11	0.08	0.223
Number previous directorships same industry	0.26	0.07	0.000
Busy	0.64	0.10	0.000
Director age	54.1	54.6	0.353
Background academic	0.021	0.000	0.001
Background finance	0.094	0.046	0.002
International work experience	0.130	0.018	0.000

4. Conclusion

- First, we find that it is possible to construct an algorithm that can predict whether a particular individual will be successful as a director in a particular firm.
- Second, we compare alternative approaches to forecasting director performance and prove that machine learning methods do better than traditional econometric approaches.
- Third, we provide evidence that director popularity is related to their expected value and that shareholder support is thus a meaningful proxy for their performance.
- Finally, we use the selections from the algorithms as benchmarks to understand the process through which directors are actually chosen and identify the types of individuals who are more likely to be chosen as directors counter to the interests of shareholders.