Three Essays on the Firm's Objective and Corporate Governance

by

Alexandr S. Moskalev

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2020

Doctoral Committee:
Associate Professor David A. Miller, Co-Chair
Associate Professor Martin C. Schmalz, Co-Chair
Professor Francine Lafontaine
Professor Uday Rajan

Alexandr S. Moskalev

moskalev@umich.edu

ORCID iD: 0000-0001-6993-028X

Dedication

I dedicate this work to the people who made it possible. I thank my mother, Galina, and my father, Sergey, for supporting me in my educational endeavors. I express gratitude to my wife, Evgeniya, for helping me to stay on a path to completion of my Ph.D. degree.

Acknowledgments

I thank my advisers David Miller, Martin Schmalz, Uday Rajan, and Francine Lafontaine; and following people for their valuable advice: Tilman Börgers, Sugato Bhattacharyya, Rosina Rodriguez Olivera, Nadya Malenko, E. Han Kim, Emil Lakkis, Virginia Traweek, participants of Econ 695/696 presentations, and participants of Ross School of Business reading group in Finance and Finance Brown Bag seminar. I am also grateful to UofM Yottabyte Research Cloud for providing computing resources.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

In my dissertation, I scrutinize the notion of profit maximization being the objective of a firm. I explore the influence of widely diversified shareholders on corporate governance through voting at director elections. I proceed in three major steps. First, in a theoretical model I show that correlation in voting behavior is associated with the weights shareholders have in determination of firm's objective. Next, I provide empirical evidence that portfolio structure matters for voting decisions, and that voting decisions translate into noticeable effects on directors' career prospects. Finally, I discover that similar portfolios are associated with similar voting behavior, and that this result extends to groups of shareholders with a sizable impact on voting tallies in director elections.

I derive the objective function of a firm with heterogeneous shareholders. In contrast to Fisher separation theorem, I drop the price-taking assumption. Therefore, shareholders have no unanimous preferences for profit maximization. I allow shareholders to act strategically by omitting the conditional sincerity assumption and by accounting for possible correlation in their votes. I derive the exact form of the objective function and provide the equilibrium existence conditions. The resulting objective function can be approximated by a weighted sum of shareholders portfolios' profit. Shareholder groups with positive within-group correlation carry greater weight.

In the second chapter, I present an evidence that mutual fund's portfolio structure matters for its voting decisions, and that director elections, the most common type of corporate elections, have delayed consequences for nominees' career prospects. In an event study of funds' mergers, I find that a merger affects the acquiring mutual fund's voting behavior. I observe higher chances of future non-nomination for directors with lower shareholder support.

This result resonates with the literature on shareholder dissent. I find that low shareholder support is also associated with a notable decrease in the length of director's tenure at a company.

In the third chapter, I analyze the role of shareholders' portfolio / ownership structure on voting participation in director elections. I find that portfolio composition matters for how mutual funds vote. Funds with more similar portfolios are more likely to cast identical votes. An increase in within-group similarity of mutual funds' portfolios leads to an increase in the number of broker "Non-Votes". Thus, highly diversified horizontal shareholding causes lower participation ("rational apathy") among other shareholders. This effect gives widely diversified cohorts of mutual funds, shareholders of the firm, a higher marginal influence at director elections than their plain share of ownership would suggest.

Chapter 1

Objective Function of a Non-Price-Taking Firm with Heterogeneous Shareholders

1.1 Introduction

Quite often, Finance and Industrial Organization literature assume that a firm has the simple objective of profit maximization. This assumption rests on a classical result, Fisher separation theorem (Fisher, 1930). It asserts that shareholders with heterogeneous portfolios unanimously agree on the objective of profit maximization if firms are price-takers (Milne, 1974; Hart, 1979; DeAngelo, 1981). While we may model a finite number of firms as price-takers under certain conditions (Fama & Laffer, 1972), these conditions are unlikely to be satisfied in the real world. This paper derives a firm's objective function without relying on the price-taking assumption.

A few notable attempts in the literature have been made to relax this assumption. Rotemberg (1984) omits the price-taking assumption but constrains the model to identical investors. He shows that equilibrium with shareholder unanimity exists under the assumption that firms act in the interest of the owners of their capital. Milne (1981) examines the firm's objective function as a collective choice problem. He establishes that in light of Arrow's Possibility Theorem (Arrow, 1963), in general, shareholders' preferences cannot be aggregated into a firm's objective function.

Grossman & Stiglitz (1977) reconcile the conventional profit maximization objective with the view of maximization of manager's utility in a dynamic setting. They conclude that unanimity cannot be assured unless "competitivity" (price-taking) assumption is satisfied. Hart (1979) and DeAngelo (1981) confirm the necessity of price-taking assumption for unanimity.

The present paper derives a micro-founded firm's objective function when firms are not perfect competitors. I start by confining the model to the one-period static case where shareholders hold exogenously defined heterogeneous portfolios. Every firm must decide on a single-dimensional policy variable from a compact convex-set. The payoff from firms' activities to shareholders is certain once all policy decisions have been made. Between-firm externalities are embedded into this deterministic payoff to shareholders. I do not assume any particular form of competition between the firms.

The unanimity condition is often relaxed in the literature. For instance, Benninga & Muller (1979) provide conditions of existence of a production equilibrium that is supported by a majority of firm's shareholders. Harris & Raviv (1988) derive conditions that make the simple majority with "one share - one vote" voting rule socially optimal. DeMarzo (1993) views firm's objective as being defined by a decision mechanism based on shareholder control. Hansen & Lott (1996) and Crès & Tvede (2005) study the conditions of sufficient internalization of between-firm externalities via majority voting among shareholders. The latter states "[a firm's] behavior is modeled as representing the shareholders' interests through a centralized political process". That is, the firm's objective is defined as adopting the production plan that cannot be defeated in the majority elections.

In this paper the objective function of a firm arises from the corporate governance process. Shareholders delegate the policy decision to an elected manager or director slate.[1] At each firm a pair of office-motivated candidates announce their policy choices that they commit to. The candidate who receives more than half of the votes wins the office. A candidate with exactly half of the votes becomes the manager with probability 1/2. Hence, the policy selection process of the winning candidate embodies the firm's objective function as his proposal gets implemented.

The value of the derived objective function goes beyond just the theoretical interest. With the rise in common ownership over the last two decades (Davis, 2008; Lindsey, 2008; Matvos & Ostrovsky, 2008; Harford et al., 2011), questions about its effects gain more traction among researchers. Common ownership is associated with lower industry returns, anticompetitive effects in airline industry, higher probability of mutual fund opposing management in elections, higher CEO compensation, and other effects (see Schmalz, 2018 for a review). Profit maximization is unlikely to be the firm's objective in the absence of perfect competition (Hansen & Lott, 1996; Gordon, 2003). The current alternative in the literature is the maximization of a weighted sum of profits of shareholders' portfolios (Salop & O'Brien, 2000). My paper contributes to the literature by establishing the sufficient conditions that allow this representation of objective function and by providing a procedure for recovering the weights for different shareholders. For instance, European Commission case regarding Dow/DuPont merger (European Commission, 2017) has failed to use Modified Herfindahl-Hirschman index of market concentration because the required control weights were not computed. My work provides a straightforward path for computing these weights starting from the voting rights of corresponding shareholders.

I model the corporate governance process with help of an *additive bias model* of manager elections (Banks & Duggan, 2005). Prospective managers compete for office in majority

[1] In the simplest meaningful case, there are two candidates for this position. For example, shareholders may voice their view through nomination of a contender manager to challenge the incumbent manager's policy.

vote elections with no abstention. They propose firms' policies in order to attract votes of shareholders. Managers do not derive any utility from the implemented policy per se, but from holding office. Shareholders vote strategically at every firm to maximize their expected utility. Each shareholder has a symmetric random bias (taste shock) that is realized after all candidates made their proposals. The bias component serves two purposes: (i) it "smoothens" otherwise discontinuous shareholder's best response function, and (ii) it allows to account for correlated behavior of different shareholders. Institutional shareholders tend to vote according to their ideology: some prefer to align their votes with management recommendations while other prefer to vote according to proxy advisors' recommendations (Bolton et al., 2018; Bubb & Catan, 2018). My model allows the correlated behavior to accommodate such party preferences of institutional shareholders.

The two most related papers in the literature are Azar (2017) and Brito et al. (2018). Both of them consider a similar corporate governance model with probabilistic voting. Here I briefly summarize the differences in my approach. First, current literature, to my knowledge, only considers the case of independent random bias components. This hinders the ability to accommodate shareholder ideology in voting. Correlated bias components allow for a greater influence of small shareholders of the same ideology on firm's behavior. Second, Brito et al. (2018) explicitly and Azar (2017) implicitly assume conditional sincerity of shareholders. That is, shareholders decide on votes at one firm, taking their decisions at other firms as given. I consider an equilibrium refinement that excludes weakly dominated strategies of shareholders (Alesina & Rosenthal, 1996). This way shareholders act strategically by taking into account their decisions at all firms simultaneously. Third, both Azar (2017) and Brito et al. (2018) establish the objective function by matching the first-order condition of manager's problem. This approach delivers the same equilibrium point, but it may not provide conclusive evidence on the form of the objective function.[2] I adopt a more

[2]That is, both papers essentially assume that objective function of a manager is linear in portfolio profits of all shareholders. This works as a linear approximation but it should not be taken for the exact objective function.

transparent approach where I explicitly derive a Taylor approximation for an exact objective function. Fourth, I find an additional condition for equilibrium existence, which is missing in both papers. The bias component must be large enough (but limited) for an equilibrium to exist.

The present work also features a geometric analysis that allows me to solve the model with correlated random bias components. I introduce a mapping of the voting model into a set of polytopes in a multidimensional space. This analysis bypasses the hurdle of a symbolic approach in deriving the final result.

In the application of geometric analysis, I rely on the results of Lasserre (1983), Khachiyan (1993), Avis et al. (1996), and many others. I'm able to improve on some of their results. In particular, I clarify the conditions of second order differentiability of polytope's volume with respect to location parameters of its hyperplanes. This is a compromise between the strong but restrictive result of Khachiyan (1993) and the weak result under weak conditions of Lasserre (1983). I also weaken conditions of continuity of polytope's volume with respect to hyperplanes' positions presented in Lasserre (1983).

1.2 Theoretical model

This chapter introduces the general version of my theoretical model. I setup the model and list the assumptions needed in order to establish equilibrium existence. A geometric interpretation is presented to simplify most proofs.

1.2.1 Setup

An industry consists of N firms that are owned by I investors. I assume that investors are *external* in the sense of Brito et al. (2018), i.e. firms do not own control or cash flow rights in other firms.[3] Investor-i is characterized by her cash flow rights vector β_i and control rights vector γ_i. This characterization is exogenous to the model. For any firm-n I have $\sum_{i=1}^{I} \beta_{in} = 1$ and $\sum_{i=1}^{I} \gamma_{in} = 1$ with $\beta_{in}, \gamma_{in} \geq 0$.

Each firm has an unidimensional policy variable, q_n, that is set by the manager in charge.[4] I assume $q_n \in [0, \bar{Q}] \subset \mathbb{R}$, where $\bar{Q} > 0$ is an upper boundary. The manager is elected by investors under simple majority rule with no abstention in competition between two prospective managers, $\{A_n, B_n\}$. The managers are *office-motivated*, risk-neutral, and have utility of 1 if elected and 0 otherwise. A manger can only hold position in one firm, there are $2N$ prospective managers in total. The choice of firm's policy does not affect manager's utility.

Investors have a two-part utility function. First part, $R_i(q; \beta_i, \gamma_i)$, is deterministic, given the firms' policies, and represents the benefit coming from owning cash flow and control rights in firms. Second part is stochastic and reflects investor's bias towards particular manager being elected. Let q be the vector of established firms' policies and m be the vector of elected managers (where $m_n \in \{A_n, B_n\}$), then investor-i's utility function is

$$u_i(q, m; \beta_i, \gamma_i) = R_i(q; \beta_i, \gamma_i) + \delta \sum_{n=1}^{N} \mathbb{1}[m_n = A_n]\xi_{in}, \tag{1.1}$$

where $\mathbb{1}[m_n = A_n]$ is an indicator function equal to 1 when manager A_n wins, $\delta > 0$ is a scale parameter, and ξ_{in} is a random utility shock associated with the election of manager A_n at firm n. For any firm n, the joint distribution of $\xi_{in}, i \in \{1, ..., N\}$ is assumed to be symmetrical around 0.[5] Function $R_i(q; \beta_i, \gamma_i)$ is two-times differentiable and strictly concave in policy of any firm, fixing the policies of other firms.

[3] This is not a significant restriction since a non-degenerate case of ownership structure involving *internal* investors can be represented by a specific control and cash flow rights structures of external investors only. See Brito et al. (2018) for details.

[4] While generalization to multidimensional policy case is possible, I stick to the unidimensional case for the sake of simplicity.

[5] The symmetry assumption is needed in order to guarantee that prospective managers are ex-ante identical in the eyes of investors if they propose the same policies for the firm.

The game has two stages. At the first stage, potential managers simultaneously propose policies in order to attract the votes of investors. Policy proposals may take form of probability distribution over feasible policies or a single policy. After that, the nature decides on the investors' taste shocks. Then, at the second stage, investors simultaneously decide whom to vote for at every firm in order to maximize their utility. Finally, at the end of the game, policy proposals of winning candidates get implemented. In a case of probabilistic policy proposals the uncertainty is resolved at the end of the game by the nature selecting a single value from a proposed probability distribution over the feasible policies.

The game proceeds as follows:

- Investors are exogenously endowed with cash flow, β_i, and control rights, γ_i, in firms. Information on β_i and γ_i for all i is public.
- Prospective managers propose their policies simultaneously at all firms. For a firm, n, these proposals are $q_n^{A_n}$ and $q_n^{B_n}$, possibly in a form of probability distribution over the policy space $[0, \bar{Q}]$.
- Random utility shocks ξ_{in} get realized for all firms, n, and investors, i.
- Investors simultaneously decide whom to vote for at every firm.
- Managers get elected and implement the proposed policies. The nature decides on a policy outcome in a case of a probabilistic proposal being used.

For an arbitrary firm n, let C be a placeholder for the candidate in question, let s^M be the strategy profile of all candidates, and let $m_n^i(s^M) \in \Delta\{A_n, B_n\}$ be the strategy of investor-i. Then the random variable $V_C = \sum_{i=1}^{I} \mathbb{1}\left[m_n^i(s^M) = C\right] \gamma_{in}$ is the share of votes candidate C receives.

Definition 1.2.1. *For a given set $\{R_i(\cdot; \beta_i, \gamma_i); \beta_i, \gamma_i\}_{i=1}^{I}$ of model parameters, a Subgame Perfect Nash equilibrium with no weakly dominated strategies is a strategy profile $\left(s^M, s^I(s^M)\right)$, where $s^M = \left((q_n^{A_n}, q_n^{B_n})\right)_{n=1}^{N}$ is the strategy profile of prospective managers, and $s^I(s^M) =$*

$\left(\left(m^i_n(s^M)\right)_{n=1}^N\right)_{i=1}^I$ *is the strategy profile of investors, such that for every firm, n, and every candidate, C, at that firm we have*

$$q_n^C \in \underset{q_n^C}{\operatorname{argmax}}\, \mathbb{E}\left[\mathbb{1}\left[V_C > \frac{1}{2}\right] + \frac{1}{2}\mathbb{1}\left[V_C = \frac{1}{2}\right] \middle| s^M_{-q_n^C}\right] \tag{1.2}$$

and for every investor, i, we have

$$m^i \in \underset{m^i \in (\Delta\{A_n, B_n\})_{n=1}^N}{\operatorname{argmax}} \mathbb{E}\left[u_i\left(q(s^M, s^I), m(s^M); \beta_i, \gamma_i\right) \middle| s^I_{-m^i}; s^M\right], \tag{1.3}$$

where

- $C \in \{A_n, B_n\}$ *is the candidate at firm n*
- q_n^C *is the strategy of candidate C at firm n*
- $s^M_{-q_n^C}$ *strategies of all candidates except candidate C*
- m^i_n *is the strategy of investor i at firm n, $m^i_n \in \Delta\{A_n, B_n\}$*
- m^i *is the strategy of investor i at all firms*
- $m(s^M)$ *is the vector of elected candidates given their strategies s^M*
- $q(s^M, s^I)$ *is the vector of firms' policies*
- *candidates who equally split votes secure position with probability $\frac{1}{2}$*
- *candidates and investors do not play weakly dominated strategies.*

That is no candidate at any firm can change his strategy in a way that increases his chances to secure the office position, and no investor can increase her utility by changing her voting behavior.[6]

[6] A special note regarding mathematical expectations in conditions 1.2 and 1.3 is needed. Condition 1.2 uses expectation to accommodate the taste shocks of investors and account for the mixed strategies of investors. In condition 1.3 expectation takes care of the mixed strategies of investors, the probabilistic strategies of prospective managers, and the tie cases when candidates split votes in half. The taste shocks have already been realized by this moment.

1.2.2 Equilibrium refinement

The rather mild assumption of excluding weakly dominated strategies allows me to simplify the backward induction analysis for investors. The following lemma summarizes the result I get for them.

Lemma 1. *If at an arbitrary firm-n prospective managers submit identical proposals,* $q_n^{A_n} = q_n^{B_n}$*, then an investor-i's choice of the manager to vote for is based on her taste shock only. That is investor-i votes for* A_n *if* $\xi_{in} > 0$ *and she votes for* B_n *when* $\xi_{in} < 0$*.*

Proof. See Appendix. □

In a case of a tie, $\xi_{in} = 0$, I assume that investor flips a fair coin to decide. The following lemma helps us to characterize the equilibrium.

Lemma 2 (Equal probability)**.** *Every candidate has exactly 50% chance of being elected in an equilibrium.*

Proof. See Appendix. □

Given the symmetry of the model with respect to prospective managers within a firm, I will pursue a symmetric equilibrium where these candidates propose identical non-random policies.[7]

Further, in order to prove equilibrium existence, we will need to know the off equilibrium path strategies of investors. The following lemma summarizes this result.

Lemma 3. *If for all firms, k, except one firm, n, the policy proposals, possibly probabilistic, are the same,* $q_k^{A_k} = q_k^{B_k}$*, and for that one firm policy proposals are different,* $q_n^{A_n} \neq q_n^{B_n}$*,*

[7] This refinement does not reduce the competitiveness among the candidates. The individual best response depends on the actions of the other candidate, managers' actions at the other firms, shareholders' payoff and bias structures.

then investor-i votes for the prospective manager A_n if

$$\mathbb{E}\left[R((q_1, ..., q_{n-1}, q_n^{A_n}, q_{n+1}, ..., q_N)^\intercal; \beta_i, \gamma_i)\right] + \delta\xi_{in} > \\ > \mathbb{E}\left[R((q_1, ..., q_{n-1}, q_n^{B_n}, q_{n+1}, ..., q_N)^\intercal; \beta_i, \gamma_i)\right], \quad (1.4)$$

where $q_k = q_k^{A_k} = q_k^{B_k}$ for $k \neq n$, and the expectation resolves uncertainty in a case of probabilistic policy proposals.

Proof. See Appendix. □

1.2.3 Taste shocks structure

The model admits possible correlation in voting decisions of shareholders. Such correlation may arise due to investor ideology (Bolton et al., 2018; Bubb & Catan, 2018), influence of proxy advising firms (Ertimur et al., 2013; Iliev & Lowry, 2015), or by following management recommendations. The correlation is imposed onto the taste shock component of investor's utility function.

Any admissible correlation structure is supported. The following assumption defines the taste shock structure.

Assumption 1. *At an arbitrary firm-n there are $M_n > 0$ independent, symmetric around 0, uniform random variables*[8]*, ζ_{nk}, with standard deviations $\sigma_1, ..., \sigma_{M_n}$. Every investor at firm-n is characterized by M_n weights, $w_1^i, ..., w_{M_n}^i$ (with at least one being different from zero). And the taste shocks for investors at firm-n are*

$$\xi_{in} = \sum_{k=1}^{M_n} w_k^i \zeta_{nk}.$$

[8] A more advanced model is feasible: if joint probability density function for $\zeta_{nk}, \forall k$ is real continuous (positively) homogeneous function within the unit cube then the derivation follows the same route and relies on the formula provided in Lasserre (1998).

This assumption greatly simplifies the solution process. It allows for a simple geometric interpretation of probability that a coalition supports the candidate. While the uniformity assumption for ζ_{nk} has an impact on the policy outcome in the equilibrium, the main effect on established polices rather comes from correlation between the taste shock variables ξ_{in}. The next few technical results provide some useful properties of the taste shocks structure.

Lemma 4 (Symmetry of taste shocks)**.** *Given assumption 1, the taste shocks have joint distribution that is symmetrical around zero in the following fashion.*

$$P\left(\{d_1\xi_{1n} \leq d_1x_1, ..., d_I\xi_{In} \leq d_Ix_I\}\right) = P(\{d_1\xi_{1n} \geq -d_1x_1, ..., d_I\xi_{In} \geq -d_Ix_I\}),$$

for $\forall(x_1, ..., x_I)$*, and* $\forall(d_1, ..., d_I) \in \times_{i=1}^{I}\{-1, 1\}$*, where* $(d_1, ..., d_I)$ *determine the direction of inequality signs.*

Proof. See Appendix. □

Lemma 5 (Taste shocks correlation)**.** *Given assumption 1, at an arbitrary firm-n, taste shocks of investors i and j are correlated in the following way*

$$\mathbb{C}orr(\xi_{in}, \xi_{jn}) = \frac{\sum_{k=1}^{M_n} w_k^i w_k^j \sigma_k^2}{\sqrt{\sum_{k=1}^{M_n}(w_k^i)^2\sigma_k^2}\sqrt{\sum_{k=1}^{M_n}(w_k^j)^2\sigma_k^2}}.$$

Proof. See Appendix. □

Lemma 5 shows that any admissible correlation structure can be achieved.

1.3 Geometric interpretation

The assumption 1 allows me to replace a rather difficult candidate's problem of winning probability maximization with a simpler geometric problem.

Consider a substitution for assumption 1's weights, $w_k^i \to \sqrt{12} w_k^i \sigma_k$[9]. It preserves full generality while allowing me to make ζ_{nk} i.i.d. uniform on $\left[-\frac{1}{2}, \frac{1}{2}\right]$. To avoid excessive notation I omit the index of the firm, n, since almost everywhere I work with an arbitrary firm-n.

The joint probability distribution of ζ_{nk}, $k \in \{1, ..., M_n\}$ can then be represented by a unit cube in M_n-dimensional space with a center at $(0, ..., 0)^\intercal$. The cube's volume represents a set of possible realizations for the vector of random variables $(\zeta_{n1}, ..., \zeta_{nM_n})$. We can assign a mass to this cube that represents the probability of this vector being within the cube. It is reasonable to normalize such mass to 1 since the probability of this event is 1. Following the same path, the cube's mass density can be described by joint probability distribution function. Since the probability that vector $(\zeta_{n1}, ..., \zeta_{nM_n})$ lands inside an arbitrary M_n-dimensional shape is the integral of joint probability function over this shape, this probability also equals to the mass of the shape given the mass density function defined above. For a uniform and independent random variables, ζ_{nk}, the mass density is uniform within the cube and zero everywhere else. This allows us to calculate the probability of $(\zeta_{n1}, ..., \zeta_{nM_n})$ being in a certain subset by evaluating the volume of this subset within the cube.

1.3.1 Shareholder's decision and coalition formation

Lemma 3 provides the condition (inequality below) for an investor-i to vote for the candidate A_n. The following transformation shows how this condition can be translated into the geometric counterpart.

$$\xi_{in} > \frac{\mathbb{E}\left[R((q_1, ..., q_n^{B_n}, ..., q_N)^\intercal; \beta_i, \gamma_i)\right] - \mathbb{E}\left[R((q_1, ..., q_n^{A_n}, ..., q_N)^\intercal; \beta_i, \gamma_i)\right]}{\delta}$$

To compactify the inequality, denote the right hand side by $\frac{\bar{R}_i(q_n^{B_n}) - \bar{R}_i(q_n^{A_n})}{\delta}$. Replacing ξ_{in} by assumption 1, while redefining the weights to make the switch to normalized ζ_{nk} random

[9]The square root of 12 appears from the variance formula for uniform distribution.

variables, I get

$$\sum_{k=1}^{M_n} \sqrt{12} w_k^i \sigma_k \zeta_{nk} > \frac{\bar{R}_i(q_n^{B_n}) - \bar{R}_i(q_n^{A_n})}{\delta}. \tag{1.5}$$

This inequality describes a half-space in $\mathbb{R}^{M_n}$ formed by a plane with a normal vector $\sqrt{12}\left(w_1^i\sigma_1, ..., w_{M_n}^i\sigma_{M_n}\right)$. The plane contains the origin point when the candidates propose identical strategies. Investor-i votes for the candidate A_n when realization of ζ_{nk}, $k \in \{1, ..., M_n\}$, lands into this half-space. A geometric way to calculate the probability of such event is to compute the volume of the convex shape formed by intersection of this half-space with the unit cube. Since the uniform distribution of $(\zeta_{n1}, ..., \zeta_{nM_n})$ does not contain any mass points I can consider a closure of the half-space defined by inequality 1.5.[10]

The most elegant contribution of this interpretation is the ability to characterize the probability of an event when a coalition of investors votes for the same candidate. Let ϑ_1 be a coalition of investors that vote for manager A_n, $\vartheta_1 = \{i | i \in \{1, ..., I\}, m_n^i = A_n\}$. To conveniently describe a coalition, consider a vector θ_1 such that $\theta_{1i} = 1$ if investor-i belongs to the coalition, and $\theta_{1i} = 0$ otherwise.

Construct a convex shape by intersecting a unit cube with a set of half-spaces defined by a modified version of the inequality 1.5,

$$(2\theta_{1i} - 1) \sum_{k=1}^{M_n} \frac{\sqrt{12} w_k^i \sigma_k \zeta_{nk}}{||\sqrt{12} w^i \odot \sigma||} \geq (2\theta_{1i} - 1) \frac{\bar{R}_i(q_n^{B_n}) - \bar{R}_i(q_n^{A_n})}{\delta ||\sqrt{12} w^i \odot \sigma||}, \tag{1.6}$$

for all $i \in \{1, ..., I\}$, where $\sigma = (\sigma_1, ..., \sigma_{M_n})$, $w^i = (w_1^i, ..., w_{M_n}^i)$, $||\sqrt{12} w^i \odot \sigma|| = \sqrt{12 \sum_{k=1}^{M_n} w_k^i \sigma_k w_k^i \sigma_k}$, and $\odot$ represents Hadamard product.[11] This shape represents the set of realizations of $(\zeta_{n1}, ..., \zeta_{nM_n})$ that correspond to voting behavior prescribed for coalition ϑ_1. The volume of the shape reflects the probability of this event.

[10] The inequality 1.5 defines a half-space that does not include the boundary formed by the hyperplane. The M_n dimensional volume of the boundary is zero, and, since uniform distribution has no mass points, the probability of random point being on a boundary is zero. Including the boundary makes the half-space closed which simplifies the further steps: intersection of several such half-spaces results in a closed convex shape.

[11] Modification consists of normalization and accounting for the "sign" to choose the correct half-space. Hadamard product for two matrices A and B of the same dimension is defined as $(A \odot B)_{i,j} = (A)_{i,j}(B)_{i,j}$.

The shape is a convex $\mathcal{H}$-polytope where halfspaces are defined by inequality 1.6 for $i \in \{1, ..., I\}$ together with the halfspaces that form the unit cube.

1.3.2 Comparative statics for a single coalition

Comparative statics has a geometric interpretation as well. The candidate A_n may want to know how an update in his policy proposal, Δq, affects the probability that coalition ϑ_1 forms and supports him. For simplicity, I consider a marginal change in policy proposal so the corresponding polytope does not change much.[12] Multiple investors may be affected by a such change.

For now, consider an effect that this change has on an arbitrary investor-i. WLOG assume that $\theta_{1i} = 1$. Then the factor in the right hand side of inequality 1.6 will be $\frac{\bar{R}_i(q_n^{B_n}) - \bar{R}_i(q_n^{A_n}) - \Delta \bar{R}_i}{\delta}$, where $\Delta \bar{R}_i = \bar{R}_i(q_n^{A_n} + \Delta q) - \bar{R}_i(q_n^{A_n})$. Hence, a policy adjustment results in a shift of hyperplane for investor-i along the normal vector $\sqrt{12}\left(w_1^i \sigma_1, ..., w_{M_n}^i \sigma_{M_n}\right)$ governed by a change in investor-i's deterministic component of utility function, $\Delta \bar{R}_i$. If this hyperplane was forming a $(M_n - 1)$-dimensional face, facet, of the polytope then the shift would result in a change of volume, and, consequently, in a change of probability that coalition ϑ_1 forms to support A_n.

To quantify the size of the change I take a derivative of the right hand side of inequality 1.6 with respect to $q_n^{A_n}$.

$$\frac{\partial RHS}{\partial q_n^{A_n}} = -\frac{(2\theta_{1i} - 1)}{\delta || \sqrt{12} w^i \odot \sigma ||} \frac{\partial \bar{R}_i(q_n^{A_n})}{\partial q_n^{A_n}} \tag{1.7}$$

This expression gives us the direction and the rate at which investor-i's hyperplane moves along the normalized normal vector $\frac{(2\theta_{1i}-1)}{||w^i \odot \sigma||}\left(w_1^i \sigma_1, ..., w_{M_n}^i \sigma_{M_n}\right)$. Note that this vector points towards the newly formed half-space. See inequality 1.6. This means that a negative value

[12] Since even a smallest change may easily change the number of vertices/faces of the polytope I need to be a bit more precise here. By "not changing much" I mean that the moving hyperplane will not cross any vertices formed by the other hyperplanes. This excludes any vertices that originally were at the moving hyperplane.

in equation 1.7 represents a movement of the hyperplane along the normal vector in the opposite direction that captures new volume into the formed half-space.

There is one more missing link between the derivative of the polytope volume and equation 1.7. Since the hyperplane defined by inequality 1.6 forms[13] only one facet of the polytope at best, its effect on polytope's volume is limited. It is proportional to the $M_n - 1$ dimensional volume of this facet and the rate of shift of the hyperplane. Using the Proposition 3.3 from Lasserre (1983) I can prove the following result.

Proposition 1. *Given a coalition ϑ_1, marginal change in probability, $\mathbb{V}(\theta_1)$, that this coalition is going to form and support prospective manager A_n solely due to actions of investor-i is*

$$\frac{\partial \mathbb{V}(\theta_1)}{\partial \bar{R}_i(q_n^{A_n})} = \frac{(2\theta_{1i} - 1)\mathbb{V}_i(\theta_1)}{\delta||\sqrt{12}w^i \odot \sigma||}, \tag{1.8}$$

where $\mathbb{V}(\theta_1)$ is volume of coalition ϑ_1's polytope, and $\mathbb{V}_i(\theta_1)$ is the $M_n - 1$ dimensional volume of facet formed by investor-i's hyperplane.

Proof. See Appendix. □

An immediate next step is to use proposition 1 to describe the effect of marginal change of policy proposal on the probability of support from coalition ϑ_1. This time I allow every investor to re-evaluate the benefit of voting for/against the prospective manager.

Corollary 1. *Given a coalition ϑ_1, marginal change in probability, $\mathbb{V}(\theta_1)$, that this coalition is going to form and support prospective manager A_n due to a change in his policy proposal, $q_n^{A_n}$, is*

$$\frac{\partial \mathbb{V}(\theta_1)}{\partial q_n^{A_n}} = \sum_{i=1}^{I} \frac{(2\theta_{1i} - 1)\mathbb{V}_i(\theta_1)}{\delta||\sqrt{12}w^i \odot \sigma||} \frac{\partial \bar{R}_i(q_n^{A_n})}{\partial q_n^{A_n}}. \tag{1.9}$$

Proof. See Appendix. □

[13] Here "forms" means that the facet also belongs to the hyperplane.

1.3.3 Candidate's problem

Prospective manager A_n may also want to know how a change in his policy proposal, $q_n^{A_n}$, affects his overall probability of getting into the office. This is a rather straightforward problem as it just requires to take a sum over all possible coalitions that assure winning for the candidate A_n.[14]

Let Θ_{A_n} be a set of coalitions at firm-n such that candidate A_n wins. Recall that γ_{in} is control rights of investor-i at firm-n. Then the vector of any coalition $\vartheta_h \in \Theta_{A_n}$ satisfies the dot product condition $\theta_h^{\intercal}\gamma_{\cdot n} \geq 0.5$. The total probability of candidate A_n winning is the sum of volumes associated with these coalitions,

$$P[m_n = A_n] = \sum_{h=1}^{\#\Theta_{A_n}} \mathbb{V}(\theta_h). \tag{1.10}$$

In equilibrium, as lemma 2 suggests, candidates split volume of the unit cube in half. Most coalitions in Θ_{A_n} have associated polytopes that share facets. Such facets are formed by investors that are non-pivotal in these coalitions. Prospective managers maximize their winning probability by catering to investors which form facets shared by both pro A_n and pro B_n coalitions. These are the pivotal investors.

In the next section we are going to exploit a great result that follows from the geometric interpretation to obtain objective function of a firm. This will allow us to establish objective function directly without matching the first order conditions as Azar (2017) and Brito et al. (2018) do. The disadvantage of matching approach is that multiple different objective functions may have the same first order conditions. The present approach removes this hurdle by directly transforming the stated manager's utility function into the firm's objective function.

[14] If a certain coalition leads to a tie between the candidates, then each candidate adds this coalition probability with half weight, i.e. tie is resolved with a coin flip. Further I implicitly assume this tie breaking rule if need arises.

1.4 Objective function

Firm's objective stems from the objective of the party that determines the firm's actions. The present work treats firm as an artificial construct, an interface, that allows involved parties (consumers, producers, managers, investors, etc.) to interact with each other in a formal way. A firm has no preferences over the outcomes apart from the ones that are implied by the decision making mechanism.

In this paper, the objective function of a firm is set by the manager who acts in his own interest conditional on restrictions put in place by the shareholders in the decision making mechanism (elections). To recover the firm's objective function I need to transform the prospective manager's utility maximization problem into the firm's optimal policy problem.

An office-motivated candidate wants to maximize the probability of winning the elections. Within the geometric interpretation this is equivalent to maximizing the volume of the unit cube that belongs to the winning coalitions. Equation 1.10 expresses this volume as a sum of volumes of polytopes associated with these.

1.4.1 The exact formulation

Without loss of generality, assume that candidate A_n will win the elections.[15] Let us fix the candidate B_n's proposal as q_n^B. For an arbitrary coalition, ϑ_1, that is winning for candidate A_n, let the volume of corresponding polytope be $\mathbb{V}(\theta_1; q_n^A, q_n^B)$. Let $\mathbb{H}$ be the set of inequalities that form the unit cube, $\mathbb{H} = \left\{(-1)^k \zeta_{n\lceil k/2 \rceil} \geq -\frac{1}{2}, \text{ for } k = 1, 2, ..., 2M_n\right\}$. For further notational convenience, let us enumerate the investor-formed hyperplanes as $\mathbb{I}(i)$ for $i = 1, ..., I$, and the walls of the unit cube as $\mathbb{H}(k)$ for $k = 1, ..., 2M_n$.

[15] Due to the symmetry of the problem the same can be done if I assume that candidate B_n will win the elections.

Using Theorem 3.1 from Lasserre (1983) I can rewrite the volume of the polytope as

$$\mathbb{V}(\theta_1; q_n^A, q_n^B) = \frac{1}{M_n}\left(\sum_{i=1}^{I} \frac{\bar{R}_i(q_n^{A_n}) - \bar{R}_i(q_n^{B_n})}{\delta||\sqrt{12}w^i \odot \sigma||}(2\theta_{1i} - 1)\mathbb{V}_{\mathbb{I}(i)}(\theta_1; q_n^A, q_n^B) + \right.$$
$$\left. + \sum_{k=1}^{2M_n} \frac{1}{2}\mathbb{V}_{\mathbb{H}(k)}(\theta_1; q_n^A, q_n^B)\right), \quad (1.11)$$

where $\mathbb{V}_{\mathbb{I}(i)}$ is the $M_n - 1$ dimensional volume of facet formed by hyperplane $\mathbb{I}(i)$, and the same is true for $\mathbb{V}_{\mathbb{H}(k)}$. The next proposition summarizes the objective function of a firm.

Proposition 2. *Assume that candidate A_n will win the elections. Then his problem of choosing the right policy proposal, against the worst possible strategy of other candidate, B_n, does essentially define the firm's objective function.*

Candidate A_n maximizes the probability of being elected,

$$\max_{q_n^A} P[m_n = A_n] = \max_{q_n^A} \sum_{h=1}^{\#\Theta_{A_n}} \mathbb{V}(\theta_h; q_n^A, q_n^B) =$$
$$= \max_{q_n^A} \frac{1}{M_n}\left(\sum_{i=1}^{I} \frac{\bar{R}_i(q_n^{A_n}) - \bar{R}_i(q_n^{B_n})}{\delta||\sqrt{12}w^i \odot \sigma||} \sum_{h=1}^{\#\Theta_{A_n}} (2\theta_{hi} - 1)\mathbb{V}_{\mathbb{I}(i)}(\theta_h; q_n^A, q_n^B) + \right.$$
$$\left. + \sum_{k=1}^{2M_n} \frac{1}{2} \sum_{h=1}^{\#\Theta_{A_n}} \mathbb{V}_{\mathbb{H}(k)}(\theta_h; q_n^A, q_n^B)\right), \quad (1.12)$$

where $\#\Theta_{A_n}$ is the set of winning coalitions, q_n^B is the worst possible (for candidate A_n) strategy of candidate B_n.

Proof. See Appendix. ∎

The advantage of proposition 2 result over the similar results in existing literature is that it is a direct result that does not come from extrapolation of manager's first order condition. Hence, it remains valid for all possible combinations of parameters and strategies no matter how far the proposed strategy is away from the equilibrium one. Yet it might be difficult to use this result further as is.

1.4.2 A Taylor approximation of the objective function

To simplify the firm's objective function I construct a Taylor approximation that is sufficient within the set of reasonable[16] strategies. Assume now that the candidate B_n plays some reasonable strategy q_n^B, while candidate A_n is playing an arbitrary strategy $\bar{q}_n^A$. This corresponds to the following payoff function for candidate A_n,

$$P[m_n = A_n] = \sum_{h=1}^{\#\Theta_{A_n}} \mathbb{V}(\theta_h; \bar{q}_n^A, q_n^B).$$

Consider candidate A_n changing his strategy to q_n^A. While the change is small, I can use Taylor expansion for individual coalitions' probabilities over the set of investors' deterministic payoff functions to evaluate candidate's A_n payoff.

I call a polytope in $\mathcal{H}$-representation *non-degenerate* if its every inequality corresponds to a half-space. I call the same polytope *non-redundant* if half-space forming hyperplanes with collinear normal vectors do not intersect.

Required differentiability is established by the following result.

Proposition 3. *Suppose that the polytope is defined in a non-degenerate and non-redundant way. Function $\mathbb{V}(\theta_h; q_n^A, q_n^B)$ is differentiable in $\bar{R}_i(q_n^A)$. Under additional condition*

$$\forall i$$

that no triplet of polytope forming half-spaces has hyperplanes, that have common point of intersection, and have linearly dependent normal vectors, the function $\mathbb{V}(\theta_h; q_n^A, q_n^B)$ is twice differentiable.

Proof. See Appendix. □

[16] I assume that a firm's strategy is not reasonable if there exists another strategy such that every investor, with positive control rights, prefers the latter one in a bilateral comparison. The set of reasonable strategies is not empty. It includes at least one point if all shareholders agree on a single strategy. If shareholders disagree, the set includes a continuum of points where shareholders disagree on the direction of change in the policy.

That is for a single coalition we have,

$$\mathbb{V}(\theta_h; q_n^A, q_n^B) = \mathbb{V}(\theta_h; \bar{q}_n^A, q_n^B) + \sum_{i=1}^{I} \frac{\partial \mathbb{V}(\theta_h; \bar{q}_n^A, q_n^B)}{\partial(\bar{R}_i(q_n^A)/\delta)} \left(\frac{\bar{R}_i(q_n^A) - \bar{R}_i(\bar{q}_n^A)}{\delta}\right) + \\ + O\left(\left|\frac{\bar{R}_i(q_n^A) - \bar{R}_i(\bar{q}_n^A)}{\delta}\right|^2\right). \quad (1.13)$$

Using the Proposition 1 , we can write down the candidate's A_n payoff as

$$P[m_n = A_n](q_n^A | q_n^B) = \sum_{h=1}^{\#\Theta_{A_n}} \mathbb{V}(\theta_h; q_n^A, q_n^B) = \sum_{h=1}^{\#\Theta_{A_n}} \mathbb{V}(\theta_h; \bar{q}_n^A, q_n^B) + \\ + \sum_{h=1}^{\#\Theta_{A_n}} \sum_{i=1}^{I} \frac{(2\theta_{hi} - 1)\mathbb{V}_i(\theta_h)}{||\sqrt{12}w^i \odot \sigma||} \left(\frac{\bar{R}_i(q_n^A) - \bar{R}_i(\bar{q}_n^A)}{\delta}\right) + O\left(\max_i \left|\frac{\bar{R}_i(q_n^A) - \bar{R}_i(\bar{q}_n^A)}{\delta}\right|^2\right), \quad (1.14)$$

where $\mathbb{V}_i(\theta_h)$ is the $M_n - 1$ dimensional volume of facet formed by investor-i's hyperplane. This equation represents an approximation of the firm's objective function when both prospective managers play the equilibrium strategy. The proposition below summarizes this result.

Proposition 4. *In an equilibrium, firm's objective function from the perspective of manager A_n can be approximated as*

$$P[m_n = A_n](q_n^A) = \sum_{i=1}^{I} \sum_{h=1}^{\#\Theta_{A_n}} \frac{(2\theta_{hi} - 1)\mathbb{V}_i(\theta_h)}{||\sqrt{12}w^i \odot \sigma||} \bar{R}_i(q_n^A). \quad (1.15)$$

Proof. See Appendix. □

That is, firm maximizes the weighted sum of shareholders' deterministic payoffs. In the base interpretation, these payoffs are their portfolios profits. The weights have two components: Banzhaf power index (Banzhaf, John F. III, 1964) and correlated voting multiplier, $\frac{\mathbb{V}_i(\theta_h)}{||\sqrt{12}w^i \odot \sigma||}$.

Absent any correlation among shareholders, $M_n = I$ and i.i.d. ξ_{in}, the correlated voting multipliers are all the same, $\frac{\mathbb{V}_i(\theta_h)}{||\sqrt{12}w^i \odot \sigma||} = \frac{\mathbb{V}_j(\theta_h)}{||\sqrt{12}w^j \odot \sigma||}$, $\forall i, j$. Then the weights, $\sum_{h=1}^{\#\Theta_{A_n}}(2\theta_{hi} - 1)$, correspond to the number of coalitions where investor-i is pivotal. That essentially translates into the Banzhaf power index up to a constant factor. This uncorrelated case effectively replicates the result in Brito et al. (2018).

When correlation is present, correlated voting multipliers reweigh the Banzhaf power index to account for joint voting behavior of shareholders. Given the complexity of possible correlation structures the exact effect on weights is hard to describe. In general, investor groups with positive correlation receive better relative representation in the objective function. In contrast, investor pairs with negative correlation tend to offset the each-others' voting power which negatively affects their relative representation. This feature of objective function is relevant as empirical evidence suggests non-trivial correlation structures in investors' votes (Bolton et al., 2018; Bubb & Catan, 2018).

1.5 Equilibrium existence

So far I learned the properties of the equilibrium assuming it exists. In this section I'm going to prove its existence. The route I'm taking is the following. I assume the properties I found earlier, derive the structure and features of best response functions, and apply a fixed-point theorem to establish existence. This closes the loop, and the assumed properties become to be derived for an existing object.

1.5.1 Concavity of candidate's objective function

Before I proceed with the proof of equilibrium existence, I would like to establish concavity of the objective function for high enough values of taste shocks scale parameter δ.

Proposition 5. *Candidate's objective function is concave in proposed policy under fixed arbitrary values of all other inputs and large enough taste shocks.*

Proof. See Appendix. □

Concavity of objective function implies that candidate's best response is a singleton. Given that both candidates are ex-ante the same and can copy each other's strategies, the maximum ex-ante probability of winning the elections is 50% in equilibrium. Hence, the peak of each objective function has to be 0.5. Since there is a single peak and the peak can be achieved by copying the opponent's strategy, peaks have to coincide in equilibrium[17], which means that both candidates propose the same policy within a firm.

1.5.2 Existence result

Building on the model elements derived above, I present the equilibrium existence result.

Proposition 6. *In the game described in sections 1.2.1, 1.2.2, and 1.2.3 with a taste shock structure compatible with conditions of proposition 3 a Subgame Perfect Nash equilibrium exists for a finite but large enough taste shocks.*

Proof. See Appendix. □

I employ sufficient conditions of proposition 3 to simplify the proof. These mildly restrict the set of feasible taste shock structures in exchange for a way to show the strict concavity of candidate's objective function.

In contrast, the magnitude requirement for the taste shocks is a necessary condition. As numerical simulation shows, prospective managers may have non-quasi-concave objective functions if the taste shocks are not large enough. In this case, they will use probabilistic policy proposals. That severely complicates the analysis and I do not study this case in my paper.

[17]This claim implicitly assumes that I condition manager's A_n objective function on the best response of manager B_n and vice versa.

1.6 Conclusion

In this paper, I derive the objective function of a non-price-taking firm with heterogeneous shareholders. In contrast to the requirements of the Fisher separation theorem, I drop the price-taking assumption as it is unlikely to be satisfied in markets with finite number of firms. To address the recent rise in common ownership, I allow for heterogeneous portfolios with arbitrary allocation of voting rights. I employ the additive bias model of managers elections to capture the corporate governance process and to derive the objective function. For further generality, I allow for possible correlation in investors' voting behavior. This provides a path for understanding the influence of proxy advisers and diversified minority shareholders on firm's objective.

My contribution to the literature consists of several parts. Within the scope of additive bias model, I find the exact form of the objective function while allowing for a correlated shareholders' voting behavior. Then, by redesigning the assumption set, I remove the need for the assumption of conditional sincerity of shareholders. I also discover that the taste shocks have to be of sufficient magnitude in order for the equilibrium to exits. This necessary condition is missing in the present literature.

I find that the firm's objective function can be approximated as a weighted sum of investors portfolios' profit. In coherence with the literature, the weights are proportional to Banzhaf power index when correlated voting is absent. These weights change drastically if shareholders are able to correlate their votes. In general, a group of shareholders with positive within-group correlations receives better relative representation in the objective function. These features of the objective function may assists the studies of influence of proxy advisers, strategic behavior of institutional investors, and the influence of common ownership on firm's behavior and competition.

Appendix A

Proof of lemma 1

An arbitrary investor-i maximizes her payoff by choosing a vector of voting decisions $m^i = (m^i_1, ..., m^i_n, ..., m^i_N)^\intercal$, where $m^i_n \in \Delta\{A_n, B_n\}$. While I can not fully describe her behavior for every possible state of input parameters (strategies of prospective managers) and all possible strategies of other investors, I can characterize her behavior with respect to candidates that propose identical policies.[18]

At firm-n candidates A_n and B_n propose $q_n^{A_n} = q_n^{B_n}$. WLOG assume that $\xi_{in} > 0$. Suppose that investor-i plays strategy m^i such that $m^i_n \neq A_n$. That is she votes for candidate B_n with some positive probability. Our goal is to prove that this strategy is weakly dominated by $\tilde{m}^i = (m^i_1, ..., A_n, ..., m^i_N)^\intercal$ if her voice is ever pivotal.

From condition 1.3 we know that investor-i maximizes expected utility conditional on actions of other players. We can rewrite that expectation in the following way using $q_n^{A_n} = q_n^{B_n}$ and equation 1.1 to substitute the utility function

$$
\begin{aligned}
&\mathbb{E}\left[u_i\left(q(s^M, s^I), m(s^M); \beta_i, \gamma_i\right)\middle| s^I_{-m^i}; s^M\right] = \\
&\qquad = \mathbb{E}\left[u_i\left(q(s^M, s^I), m(s^M); \beta_i, \gamma_i\right)\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A = q_n^B = q_n\right] = \\
&\qquad = \mathbb{E}\left[R_i(q; \beta_i, \gamma_i) + \delta\sum_{k=1}^{N}\mathbb{1}[m_k = A_k]\xi_{ik}\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A = q_n^B = q_n\right] = \\
&\qquad = \mathbb{E}\left[R_i(q; \beta_i, \gamma_i) + \delta\sum_{k=1, k\neq n}^{N}\mathbb{1}[m_k = A_k]\xi_{ik}\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A = q_n^B = q_n\right] + \\
&\qquad\qquad + \mathbb{E}\left[\delta\mathbb{1}[m_n = A_n]\xi_{in}\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A = q_n^B = q_n\right].
\end{aligned}
$$

[18] Candidates may use probabilistic strategies in their policy proposals.

Notice that the first term does not depend on action of investor-i at firm-n, m_n^i, since the implemented policy q_n, possibly probabilistic, is the same no matter which candidate wins. The second term only depends on the taste shock of investor-i at firm-n.

Several cases are possible. First, consider a situation where her voice is never pivotal, e.g. there is another investor with more than 50% of voting power. Then her vote at the firm n does not matter and she may as well vote $\tilde{m}_n^i = A_n$. And her strategy, $\tilde{m}^i$, is not weakly dominated by another one that differs only in m_n^i.

Second, she might be the controlling investor with more that 50% of voting power at firm-n. Since candidates propose the same policy, $q_n^{A_n} = q_n^{B_n}$, the first term is the same for both candidates in expectation over elected managers and implemented policies. The only difference comes from the second term, $\delta \mathbb{1}[m_n = A_n]\xi_{in}$, where we can drop the expectation since she is the controlling investor. Given $\xi_{in} > 0$, she derives higher utility from candidate A_n being the manager. Hence, her strategy is $\tilde{m}^i$, and $\tilde{m}^i$ dominates m^i.

Third, she might be pivotal sometimes depending on the actions of other investors and the nature. While both strategies m^i and $\tilde{m}^i$ bring her the same expected payoff if her voice is not pivotal, strategy $\tilde{m}^i$ brings her higher expected payoff than m^i if her voice is pivotal.[19] Hence, strategy $\tilde{m}^i$ weakly dominates m^i. By the no weakly dominated strategies assumption she will not play m^i. That is she plays $\tilde{m}^i$, and $\tilde{m}_n^i = A_n$.

Finally, we conclude that she votes for manager A_n if $\xi_{in} > 0$ and for manager B_n if $\xi_{in} < 0$. The last step is to consider $\xi_{in} = 0$. With $\xi_{in} = 0$ both candidates deliver the same expected payoff to the investor-i. Then she may vote for anyone at firm-n without perturbing her utility function.

□

Proof of lemma 2

For a pair of prospective managers at an arbitrary firm n consider the following cases:

[19] Note that as above, the expectation of the first term is the same and the difference comes from the taste shock depending on the manager elected.

- Strategies $q_n^{A_n}$ and $q_n^{B_n}$ are played such that $\mathbb{E}[\mathbb{1}[m_n = A_n]] = 0.5$. Then lemma 2 is trivially true.
- Strategies $q_n^{A_n} \neq q_n^{B_n}$ are played such that $\mathbb{E}[\mathbb{1}[m_n = A_n]] \neq 0.5$. WLOG assume that $\mathbb{E}[\mathbb{1}[m_n = A_n]] < 0.5$. Then candidate A_n can increase the winning probability by deviating to the strategy $q_n^{A_n} = q_n^{B_n}$, so we are not in equilibrium.

Candidates with identical strategies can only be distinguished by the random utility component they derive to shareholders. Since random variables ξ_{in} have symmetric joint distribution around zero the ex-ante chance to elect a certain candidate is 50%.[20] The following steps show this result.

Consider an investor-i's decision to vote for manager A_n, denote this event as $\mathbb{1}[i$ votes for $A_n]$. By lemma 1, investor-i votes for the candidate A_n if $\xi_{in} > 0$ or with probability 1/2 if $\xi_{in} = 0$.

$$\mathbb{1}[i \text{ votes for } A_n] = \mathbb{1}[\xi_{in} > 0] + x\mathbb{1}[\xi_{in} = 0],$$

where $x \sim Bernoulli(1/2)$ and is drawn independently at any tie. Recall that V_{A_n} is the share of votes submitted in the favor of the prospective manager by shareholders. Then due to the symmetry of joint distribution

$$\begin{aligned}
P\left[V_{A_n} > \frac{1}{2}\right] = P\left[\sum_{i=1}^{I}\left(\mathbb{1}[\xi_{in} > 0] + x\mathbb{1}[\xi_{in} = 0]\right)\gamma_{in} > \frac{1}{2}\right] = \\
= P\left[\sum_{i=1}^{I}\left(\mathbb{1}[-\xi_{in} > 0] + x\mathbb{1}[\xi_{in} = 0]\right)\gamma_{in} > \frac{1}{2}\right] = \\
= P\left[\sum_{i=1}^{I}\left(\mathbb{1}[\xi_{in} < 0] + x\mathbb{1}[\xi_{in} = 0]\right)\gamma_{in} > \frac{1}{2}\right] = \\
= P\left[\sum_{i=1}^{I}\left(\mathbb{1}[\xi_{in} < 0] + (1-x)\mathbb{1}[\xi_{in} = 0]\right)\gamma_{in} > \frac{1}{2}\right] = \\
= P\left[1 - \sum_{i=1}^{I}\left(\mathbb{1}[\xi_{in} > 0] + x\mathbb{1}[\xi_{in} = 0]\right)\gamma_{in} > \frac{1}{2}\right] = P\left[V_{A_n} < \frac{1}{2}\right].
\end{aligned}$$

[20] This implies that the case with identical strategies, $q_n^{A_n} = q_n^{B_n}$, but $\mathbb{E}[\mathbb{1}[m_n = A_n]] \neq 0.5$ is not possible.

Note that $P[m_n = A_n] + P[m_n = B_n] = 1$, hence

$$P[m_n = A_n] = \mathbb{E}\left[\mathbb{1}\left[V_{A_n} > \frac{1}{2}\right] + \frac{1}{2}\mathbb{1}\left[V_{A_n} = \frac{1}{2}\right]\right] =$$
$$= \mathbb{E}\left[\mathbb{1}\left[V_{A_n} < \frac{1}{2}\right] + \frac{1}{2}\mathbb{1}\left[V_{A_n} = \frac{1}{2}\right]\right] = P[m_n = B_n] = \frac{1}{2}.$$

□

Proof of lemma 3

By lemma 1 investor-i, when making her decisions at firms $k \neq n$, is going to adhere to the taste shocks only. This result holds for all investors. Hence, the voting decisions and election outcomes at other firms are not affected by the behavior of candidates at the firm-n.

The expected utility of investor-i has the following form

$$\begin{aligned}
&\mathbb{E}\left[u_i\left(q(s^M, s^I), m(s^M); \beta_i, \gamma_i\right)\middle| s^I_{-m^i}; s^M\right] = \\
&= \mathbb{E}\left[u_i\left(q(s^M, s^I), m(s^M); \beta_i, \gamma_i\right)\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A \neq q_n^B\right] = \\
&= \mathbb{E}\left[R_i(q; \beta_i, \gamma_i) + \delta \sum_{k=1}^{N} \mathbb{1}[m_k = A_k]\xi_{ik}\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A \neq q_n^B\right] = \\
&= \mathbb{E}\left[\delta \sum_{k=1, k\neq n}^{N} \mathbb{1}[m_k = A_k]\xi_{ik}\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A \neq q_n^B\right] + \\
&\quad + \mathbb{E}\left[R_i(q; \beta_i, \gamma_i) + \delta\mathbb{1}[m_n = A_n]\xi_{in}\middle| s^I_{-m^i}; s^M_{-(q_n^A, q_n^B)}, q_n^A \neq q_n^B\right],
\end{aligned}$$

where $s^M_{-(q_n^A, q_n^B)} = (q_1^{A_1}, ..., q_{n-1}^{A_{n-1}}, q_{n+1}^{A_{n+1}}, q_N^{A_N})^\intercal = (q_1^{B_1}, ..., q_{n-1}^{B_{n-1}}, q_{n+1}^{B_{n+1}}, q_N^{B_N})^\intercal$. Note that the first term does not depend on the voting strategy or on the election result at the firm-n. This allows us to concentrate on the second term only. Since at every other firm the policy proposals are the same, the implemented policies are the same as well. For the firm-n the implemented policy depends on result of the elections. The second term can be written in the following way depending on the election outcome:

$$\begin{aligned} &\mathbb{E}\left[R((q_1, ..., q_{n-1}, q_n^{A_n}, q_{n+1}, ..., q_N)^\intercal; \beta_i, \gamma_i)\right] + \delta\xi_{in}, && \text{if candidate } A_n \text{ is elected} \\ &\mathbb{E}\left[R((q_1, ..., q_{n-1}, q_n^{B_n}, q_{n+1}, ..., q_N)^\intercal; \beta_i, \gamma_i)\right], && \text{if candidate } B_n \text{ is elected} \end{aligned}$$

where the expectation remains to account for possibly stochastic policies. This result does not depend on voting strategies at firms other than firm-n. Hence, investor-i only cares about how her vote may affect the chances of candidate A_n given the strategies of other investors at the firm-n.

As in the proof of lemma 1 several cases are possible. WLOG assume that

$$\mathbb{E}\left[R((..., q_n^{A_n}, ...)^\intercal; \beta_i, \gamma_i)\right] + \delta\xi_{in} > \mathbb{E}\left[R((..., q_n^{B_n}, ...)^\intercal; \beta_i, \gamma_i)\right].$$

First, her voice may be never pivotal. Here I assume that she votes for the candidate A_n whose policy gives her higher expected utility. This assumption is optimal since her voice does not matter anyway. Second, her voice may be pivotal sometimes. Then by the exclusion of weakly dominated strategies she votes for A_n because this is not a weakly dominated strategy. If she voted for B_n with positive probability then her voice being pivotal sometimes would reduce her expected utility. Third, her voice might be always pivotal. Then she again votes for A_n since that maximizes her expected utility.

□

Proof of lemma 4

By assumption 1, ζ_{nk} is symmetric around 0, that is $-\zeta_{nk}$ is distributed the same as ζ_{nk}. Due to the linear structure of ξ_{in} and independence of ζ_{nk}, we can conclude that $-\xi_{in}$ has distribution identical to the distribution of ξ_{in}.

$$-\xi_{in} = \sum_{k=1}^{M_n} w_k^i(-\zeta_{nk}) \sim \sum_{k=1}^{M_n} w_k^i(\zeta_{nk}) = \xi_{in}.$$

Now we can generalize this result to the case of joint distribution.

$$P\left(\{d_1\xi_{1n} \leq d_1x_1, ..., d_I\xi_{In} \leq d_Ix_I\}\right) = P\left(\{-d_1\xi_{1n} \leq d_1x_1, ..., -d_I\xi_{In} \leq d_Ix_I\}\right) =$$

$$= P(\{d_1\xi_{1n} \geq -d_1x_1, ..., d_I\xi_{In} \geq -d_Ix_I\}),$$

where the first equality is justified by replacing all ζ_{nk} with $-\zeta_{nk}$ within all ξ_{in} as demonstrated above.

□

Proof of lemma 5

Consider covariance between ξ_{in} and ξ_{jn}

$$\mathbb{C}ov(\xi_{in}, \xi_{jn}) = \mathbb{C}ov\left(\sum_{k=1}^{M_n} w_k^i\zeta_{nk}, \sum_{k=1}^{M_n} w_k^j\zeta_{nk}\right) = \sum_{k=1}^{M_n} \mathbb{C}ov(w_k^i\zeta_{nk}, w_k^j\zeta_{nk}),$$

since $\mathbb{C}ov(\zeta_{nk}, \zeta_{nl}) = 0$ for $k \neq l$. Then

$$\mathbb{C}ov(\xi_{in}, \xi_{jn}) = \sum_{k=1}^{M_n} w_k^i w_k^j \mathbb{C}ov(\zeta_{nk}, \zeta_{nk}) = \sum_{k=1}^{M_n} w_k^i w_k^j \sigma_k^2.$$

Last step is to construct correlation using the above formula for covariance.

□

Proof of proposition 1

Lasserre (1983) defines polytope with the set of inequalities $a_i x \leq b_i$ and then shows that

$$\frac{\partial \mathbb{V}(n, A, b)}{\partial b_i} = \frac{\mathbb{V}_i(n-1, A, b)}{||a_i||},$$

where $\mathbb{V}(n, A, b)$ is the volume of polytope, and $\mathbb{V}_i(n-1, A, b)$ is the $n-1$ dimensional volume of facet formed by inequality $a_i x \leq b_i$. Adapting the parameters a_i and b_i to the present work I get

$$a_i = -\frac{(2\theta_{1i}-1)}{||w^i \odot \sigma||}\left(w_1^i\sigma_1, ..., w_{M_n}^i\sigma_{M_n}\right) \tag{1.16}$$

$$b_i = -(2\theta_{1i}-1)\frac{\bar{R}_i(q_n^{B_n}) - \bar{R}_i(q_n^{A_n})}{\delta||\sqrt{12}w^i \odot \sigma||}. \tag{1.17}$$

Observe that $||a_i|| = 1$, then using a chain rule I conclude

$$\frac{\partial \mathbb{V}(n,A,b)}{\partial b_i}\frac{\partial b_i}{\partial \bar{R}_i(q_n^{A_n})} = \frac{\mathbb{V}_i(n-1,A,b)}{||a_i||}\frac{2\theta_{1i}-1}{\delta||\sqrt{12}w^i \odot \sigma||} = \frac{(2\theta_{1i}-1)\mathbb{V}_i(n-1,A,b)}{\delta||\sqrt{12}w^i \odot \sigma||}.$$

$\square$

Proof of corollary 1

Recall that polytope's volume $\mathbb{V}(\theta_1)$ for coalition ϑ_1 depends on the positions of hyperplanes determined by normal vectors and offset coefficients. See inequality 1.6. These offset coefficients (the right hand side of inequality 1.6) in turn depend on $\bar{R}_i(q_n^{A_n})$. Hence, I can represent the volume of the polytope as a function that depends on $\bar{R}_i(q_n^{A_n})$ and other variables.

$$\mathbb{V}(\theta_1) = \mathbb{V}(\theta_1, \bar{R}_1(q_n^{A_n}), ..., \bar{R}_I(q_n^{A_n}), \bar{R}_1(q_n^{B_n}), ..., \bar{R}_I(q_n^{B_n}), w^1, ..., w^I, \sigma_1, ..., \sigma_{M_n})$$

Then an application of chain rule gives us

$$\frac{\partial \mathbb{V}(\theta_1)}{\partial q_n^{A_n}} = \sum_{i=1}^{I}\frac{\partial \mathbb{V}(\theta_1)}{\partial \bar{R}_i(q_n^{A_n})}\frac{\partial \bar{R}_i(q_n^{A_n})}{\partial q_n^{A_n}}.$$

The last step is to apply the proposition 1.

$\square$

Proof of proposition 2

Using equations 1.10 and 1.11 we find the total probability for candidate A_n to be elected. Since this is a constant sum game, each candidate chooses a strategy that is a best response to the worst possible move of the other candidate. Firm's objective function is, essentially, the candidate's A_n objective function that is given by equation 1.12.

$\square$

Proof of proposition 3

To prove this proposition I'm going to adopt and modify the proof of proposition 3.3 from Lasserre (1983). For this proof only, to maintain compatibility of notation I will change the notation to match the one used by Lasserre (1983). That is, the polytope's volume, $\mathbb{V}(\theta_h; q_n^A, q_n^B)$, is now $V(M_n, A, b)$, where M_n is the dimensionality of the space, and A, b (see equations 1.16, 1.17) describe the polytope as intersection of halfspaces, $\{x|Ax \leq b\}$. Recall that we have two sets of halfspaces here: ones that are formed by investors and the walls of the unit cube. Since the cube is static, we are only interested in the behavior of investors-formed hyperplanes. These are described by the inequality 1.6. Please also note that we require matrix A to have at least one non-zero element in every row. Thus we ensure non-degeneracy of the polytope representation, i.e. every inequality represents a half-space in M_n dimensional space.

First order differentiability

In the first half of the proof I will show that with no additional restrictions function $V(M_n, A, b)$ is differentiable in $\bar{R}_i(q_n^A)$. We start by establishing the form of the derivative $\frac{\partial V(M_n,A,b)}{\partial \bar{R}_i(q_n^A)}$. As inequality 1.6 suggests, the derivative of interest is proportional to the derivative $\frac{\partial V(M_n,A,b)}{\partial b_i}$, where i is the index of inequality in (A, b) representation that corresponds to the inequality 1.6 for i-th investor.

Two cases are now possible. First, the $M_n - 1$ dimensional volume of intersection of investor's i-th hyperplane with the polytope, $V_i(M_n - 1, A, b)$, is non zero. Then the following result is established by the aforementioned proposition 3.3.

$$\frac{\partial V(M_n, A, b)}{\partial b_i} = \frac{1}{||a_i||} V_i(M_n - 1, A, b), \tag{1.18}$$

where a_i is the i-th row of matrix A.

Second, the $M_n - 1$ dimensional volume of the hyperplane intersection is zero, $V_i(M_n - 1, A, b) = 0$. Here I repeat the steps from Lasserre (1983) with some modifications.

Consider an i-th hyperplane (WLOG assume that it corresponds to i-th investor, and this hyperplane is described by i-th row of matrix A) such that $V_i(M_n - 1, A, b) = 0$. For a small change, $\delta b_i > 0$ in b_i, I investigate the change in the polytope's volume

$$V(M_n, A, b + \delta b_i e_i) - V(M_n, A, b), \tag{1.19}$$

where e_i is vector of zeros everywhere except the i-th cell which contains 1.

Observe that the hyperplane i at its original place (b_i) separates space into two regions such that polytope completely resides within only one of them. The contrary violates strict unimodality of $V_i(M_n - 1, A, b)$ established by Avis et al. (1996) since sectional area can not be negative. With $\delta b_i > 0$ the hyperplane moves away from the polytope towards the region which does not contain any polytope's volume. Hence, the volume of the polytope is not changed for $\delta b_i > 0$.

Now let the small change be negative, $\delta b_i < 0$. Then the domain of the change in polytope, $\Delta(\delta b_i)$, can be described by the following set of inequalities

$$\begin{cases} b_i + \delta b_i \leq a_i \cdot x \leq b_i, \\ a_j \cdot x \leq b_j, \, \forall j \neq i. \end{cases} \tag{1.20}$$

Define new scalar variables y_k and Z as follows

$$x = x_0 + \sum_{k=1}^{M_n-1} y_k v_k + Z\frac{a_i}{||a_i||}, \tag{1.21}$$

where $a_i \cdot x_0 = b_i$, and vectors v_k, for $k = 1, ..., M_n - 1$, form an orthonormal basis of the subspace $a_i \cdot x = 0$. This allows us to represent $\Delta(\delta b_i)$ as

$$\begin{cases} \delta b_i \leq Z||a_i|| \leq 0, \\ \sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0 - Z\frac{a_j \cdot a_i}{||a_i||}, \, \forall j \neq i. \end{cases} \tag{1.22}$$

The next step is to confine the boundaries of $\Delta(\delta b_i)$ from the inside[21] and the outside. For this I introduce variables s_j and s'_j

$$s_j = \max\left(0, \delta b_i \frac{a_j \cdot a_i}{||a_i||^2}\right)$$

$$s'_j = \max\left(0, -\delta b_i \frac{a_j \cdot a_i}{||a_i||^2}\right),$$

and construct the corresponding domains $\Delta^2(\delta b_i)$ and $\Delta^1(\delta b_i)$ as follows

$$\Delta^1(\delta b_i) \text{ is the domain } \begin{cases} \delta b_i \leq Z||a_i|| \leq 0, \\ \sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0 + s'_j, \, \forall j \neq i, \end{cases}$$

and

$$\Delta^2(\delta b_i) \text{ is the domain } \begin{cases} \delta b_i \leq Z||a_i|| \leq 0, \\ \sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0 - s_j, \, \forall j \neq i. \end{cases}$$

[21]While there is no need to confine these from the inside, I do this to maintain compatibility with the proof presented by Lasserre (1983).

Now observe that the following relation holds

$$\Delta^2(\delta b_i) \subseteq \Delta(\delta b_i) \subseteq \Delta^1(\delta b_i). \tag{1.23}$$

Denote an M_n-dimensional volume function of a domain as $\mathcal{V}_{M_n}(\cdot)$, then the relation above establishes

$$\mathcal{V}_{M_n}(\Delta^2(\delta b_i)) \leq \mathcal{V}_{M_n}(\Delta(\delta b_i)) \leq \mathcal{V}_{M_n}(\Delta^1(\delta b_i)). \tag{1.24}$$

Note that for the domains $\Delta^1(\delta b_i)$, $\Delta^2(\delta b_i)$ the variable Z spans the dimension that is orthogonal to the subspace $a_i \cdot x = 0$. Hence, these domains are prismatoidal polytopes, and their volumes can be found as

$$\mathcal{V}_{M_n}(\Delta^1(\delta b_i)) = \left|\frac{\delta b_i}{||a_i||}\right| \mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) \tag{1.25}$$

$$\mathcal{V}_{M_n}(\Delta^2(\delta b_i)) = \left|\frac{\delta b_i}{||a_i||}\right| \mathcal{V}_{M_n-1}(\Delta'^2(\delta b_i)), \tag{1.26}$$

where domains $\Delta'^1(\delta b_i)$ and $\Delta'^2(\delta b_i)$ are defined as

$$\Delta'^1(\delta b_i) \text{ is the domain } \left\{ \textstyle\sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0 + s'_j, \ \forall j \neq i \right. \tag{1.27}$$

$$\Delta'^2(\delta b_i) \text{ is the domain } \left\{ \textstyle\sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0 - s_j, \ \forall j \neq i. \right. \tag{1.28}$$

Recall that at $\delta b_i = 0$ the domains above both coincide with

$$\sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0, \ \forall j \neq i, \tag{1.29}$$

which is the face of original polytope formed by the i-th investor's hyperplane. By assumption above, the $M_n - 1$ dimensional volume of this face is zero, $V_i(M_n - 1, A, b) = 0$. Since the set of inequalities 1.28 is more restrictive than the set 1.29, and the latter one has zero volume, I conclude that $\mathcal{V}_{M_n-1}(\Delta'^2(\delta b_i)) = 0$.

Next, I show that $\mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) \to 0$ as $\delta b_i \to 0$. Note that $V_i(M_n - 1, A, b) = \mathcal{V}_{M_n-1}(\Delta'^1(0))$, where $\Delta'^1(0)$ is the domain defined by 1.29. By assumption above, $V_i(M_n - 1, A, b)$ is zero, and from inequalities 1.27 and 1.29 we see that $\Delta'^1(0) \subseteq \Delta'^1(\delta b_i)$. Hence, if $\mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) > 0$, all the volume must be contained in the domain $\Delta'^1(\delta b_i)\backslash\Delta'^1(0)$. Any point from within this domain must have to violate at least one of the inequalities 1.29. Consider domains E_u, where $u \in \{1, ..., I\}\backslash\{i\}$, defined as

$$\begin{cases} \sum_{k=1}^{M_n-1} y_k(a_j \cdot v_k) \leq b_j - a_j x_0 + s'_j, \ \forall j \neq i \\ \sum_{k=1}^{M_n-1} y_k(a_u \cdot v_k) \geq b_u - a_u x_0. \end{cases} \tag{1.30}$$

The finite union of these domains covers the domain of the interest,

$$\Delta'^1(\delta b_i)\backslash\Delta'^1(0) \subseteq \cup_{u=1,u\neq i}^{I} E_u. \tag{1.31}$$

That is to show that $\mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) \to 0$ as $\delta b_i \to 0$ we need to establish $\mathcal{V}_{M_n-1}(E_u) \to 0 \ \forall u$ as $\delta b_i \to 0$.

Observe that E_u is a compact convex polytope, so it is Jordan measurable (see Tao (2011)). Tao (2010) in Proposition 1.7.2 establishes that a measurable subset E of R^{M_n-1} has a positive measure iff for any $\epsilon > 0$ there exists a ball B such that $m(E \cap B) \geq (1-\epsilon)m(B)$.

For an arbitrary u suppose that $\mathcal{V}_{M_n-1}(E_u)$ does not converge to 0 as $\delta b_i \to 0$. That is, suppose $\exists \eta > 0$, s.t. $\forall \delta b_i$ the volume $\mathcal{V}_{M_n-1}(E_u) \geq \eta > 0$[22]. Then take an arbitrary but small enough, $\frac{1}{2} > \epsilon > 0$, ϵ-ball that has its center inside E_u. Note that the ϵ-ball must have its center at a distance from every hyperplane that forms the polytope. Otherwise, the hyperplane that it lies on will dissect away half of its volume, that is $m(E_u \cap B) \leq \frac{1}{2}m(B) < (1-\epsilon)m(B)$, which violates the Proposition 1.7.2. Let the distance from the u-th hyperplane be $\mu > 0$, and realize that $s'_u \geq \mu$, since otherwise the center lies outside the polytope and

[22] Since $\mathcal{V}_{M_n-1}(E_u)$ is monotonous in δb_i and is bounded below by zero, the only possibility for $\lim_{\delta b_i \to 0} \mathcal{V}_{M_n-1}(E_u) \neq 0$ is to have $\lim_{\delta b_i \to 0} \mathcal{V}_{M_n-1}(E_u) > 0$.

Proposition 1.7.2 is violated. But s'_u is directly affected by δb_i and for any fixed $\mu(\epsilon)$ the s'_u can be made smaller than $\mu(\epsilon)$ by taking δb_i closer to zero. Hence, center of an arbitrary ϵ-ball can be excluded from the polytope for sufficiently small δb_i, which by Proposition 1.7.2 means that $\mathcal{V}_{M_n-1}(E_u) = 0$ when $\delta b_i \to 0$. Finally, it implies that

$$0 \leq \mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) = \mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i) \backslash \Delta'^1(0)) \leq \mathcal{V}_{M_n-1}(\cup_{u=1,u\neq i}^{I} E_u) \leq$$
$$\leq \sum_{u=1,u\neq i}^{I} \mathcal{V}_{M_n-1}(E_u) \to 0, \text{ as } \delta b_i \to 0. \quad (1.32)$$

And we arrive at conclusion that

$$\mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) \to V_i(M_n-1, A, b) = 0,$$

$$\mathcal{V}_{M_n-1}(\Delta'^2(\delta b_i)) \to V_i(M_n-1, A, b) = 0,$$

and

$$\lim_{\delta b_i \to -0} \frac{V(M_n, A, b+\delta b_i e_i) - V(M_n, A, b)}{\delta b_i} \leq \lim_{\delta b_i \to -0} \frac{\mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i))}{||a_i||},$$
$$\lim_{\delta b_i \to -0} \frac{V(M_n, A, b+\delta b_i e_i) - V(M_n, A, b)}{\delta b_i} \geq \lim_{\delta b_i \to -0} \frac{\mathcal{V}_{M_n-1}(\Delta'^2(\delta b_i))}{||a_i||}.$$

For the case of $\delta b_i \to +0$ we directly established above that

$$\lim_{\delta b_i \to +0} \frac{V(M_n, A, b+\delta b_i e_i) - V(M_n, A, b)}{\delta b_i} = 0. \quad (1.33)$$

This concludes our second step and we are left to state that

$$\frac{\partial V(M_n, A, b)}{\partial b_i} = \frac{V_i(M_n-1, A, b)}{||a_i||} \quad (1.34)$$

holds for all values of $V_i(M_n-1, A, b) \geq 0$.

The next step is to show that $V_i(M_n - 1, A, b)$ is a continuous function of b. Lasserre (1983) in Proposition 3.2. shows that $V(M_n, A, b)$ is a continuous function of b wherever we have $V(M_n, A, b) \neq 0$. Our goal differs in two ways: first, we need to show continuity without the restriction that volume must be positive. Second, we need to do it for a case of $M_n - 1$ dimensional polytope's volume, given that hyperplanes are specified in M_n dimensional space.

To extend Proposition 3.2 I show that $V(M_n, A, b)$ is continuous in b when $V(M_n, A, b) = 0$. Consider an arbitrary sequence c^k, such that $c^k \to b$, as $k \to \infty$. Let $D(b) = \{x | Ax \leq b\}$ be an arbitrary domain with M_n-volume being zero, $\mathcal{V}_{M_n}(D(b)) = V(M_n, A, b) = 0$. Let $D(c^k)$ be the sequence of domains formed by an arbitrary sequence of perturbations, c^k, $D(c^k) = \{x | Ax \leq c^k\}$. For an arbitrary large K two options are possible. First, for all $k \geq K$, $\mathcal{V}_{M_n}(D(c^k)) = 0$. Obviously, this implies that $\lim_{k\to\infty} \mathcal{V}_{M_n}(D(c^k)) = \lim_{k\to\infty} V(M_n, A, c^k) = 0$. Second option is that $\forall K, \exists k > K$ s.t. $\mathcal{V}_{M_n}(D(c^k)) > 0$. Observe, that this is equivalent to possibility of fitting a sufficiently small ϵ-ball inside of the polytope. This case is covered by the same logic as the proof of $\mathcal{V}_{M_n-1}(\Delta'^1(\delta b_i)) \to 0$ as $\delta b_i \to 0$ above. Hence, $\lim_{k\to\infty} \mathcal{V}_{M_n}(D(c^k)) = 0$ and this concludes the proof that $V(M_n, A, b)$ is continuous in b when $V(M_n, A, b) = 0$. Since the volume function can not be negative, this result in combination with Proposition 3.2 establishes continuity of volume function of a polytope in M_n-dimensional space[23].

Now we need to derive conditions under which a polytope defined by intersection of polytope in M_n dimensional space with a hyperplane has a continuous in b volume function.

Above we established that a polytope specified as intersection of half-spaces in R^{M_n} has continuous M_n dimensional volume with respect to changes in vector b. We can build upon this result by showing that a polytope formed by intersection of half-spaces and a hyperplane in R^{M_n} can be represented as intersection of half-spaces in R^{M_n-1}. Since this polytope lies

[23] Berger (1987) establishes continuity of volume function on the set compact convex sets with Hausdorff metric. Lasserre (1983) provides a proof for a particular case of convex polytopes. The present work extends his result by relaxing unnecessary assumptions.

within the hyperplane, we need to exclude the dimension orthogonal to the hyperplane. A similar procedure has been done with the help of equation 1.21.

Let the domain of the polytope be described as $D(A, b) = \{x \in R^{M_n} | Ax \leq b\}$. Let i be the index of equation for the hyperplane of interest. Then a following substitution for x allows us to eliminate the dimension orthogonal to hyperplane i.

$$x = x_0 + \sum_{k=1}^{M_n - 1} y_k v_k + Z \frac{a_i}{||a_i||}, \tag{1.21}$$

where $a_i \cdot x_0 = b_i$, Z is a scalar, and vectors v_k, for $k = 1, ..., M_n - 1$, form an orthonormal basis of the subspace $a_i \cdot x = 0$. This substitution gives us the following form for the polytope formed by intersection of polytope $D(A, b)$ and the i-th hyperplane.

$$\begin{cases} 0 \leq -Z \frac{a_i \cdot a_i}{||a_i||} \\ \sum_{k=1}^{M_n - 1} y_k (a_j \cdot v_k) \leq b_j - a_j \cdot x_0 - Z \frac{a_i \cdot a_j}{||a_i||}, \forall j \neq i, \end{cases} \tag{1.35}$$

where $Z = 0$, and the first inequality coming from the i-th hyperplane is redundant. The set of inequalities reduces to

$$\left\{ \sum_{k=1}^{M_n - 1} y_k (a_j \cdot v_k) \leq b_j - a_j \cdot x_0, \forall j \neq i. \right. \tag{1.36}$$

The $M_n - 1$ dimensional volume of this polytope is a continuous function of right hand side of inequalities if every inequality corresponds to a half-space in $\mathbb{R}^{M_n - 1}$. This condition can be written as $\forall j \neq i$, $\exists k$ s.t. $a_j \cdot v_k \neq 0$. That is for every inequality we can find a normal vector to the hyperplane that separates half-spaces. If for an inequality j, $\forall k$ $a_j \cdot v_k = 0$, then separating hyperplane is not defined, and the inequality works as an indicator function. This may create a discontinuity when inequality becomes pivotal.

This concludes our proof of first order differentiability.

Second order differentiability

The second order differentiability is essentially the first order differentiability of first order derivatives. Since first order derivatives rely on computation of $M_n - 1$ dimensional volume of polytope formed by cross section with a hyperplane, we can recursively apply our arguments from the first half of the proof.

Given the form of the first order derivative,

$$\frac{\partial V(M_n, A, b)}{\partial b_i} = \frac{V_i(M_n - 1, A, b)}{||a_i||}, \tag{1.34}$$

we conclude that the second order derivative has the following form

$$\frac{\partial^2 V(M_n, A, b)}{\partial b_j \partial b_i} = \frac{1}{||a_i||} \frac{V_{ij}(M_n - 2, A, b)}{||(a_j \cdot v_1, ..., a_j \cdot v_{M_n - 1})||}, \text{ for } j \neq i. \tag{1.37}$$

This formula uses definitions introduced in equation 1.36. Here we also use notation $V_{ij}(M_n - 2, A, b)$ to denote the $M_n - 2$ dimensional volume of polytope formed by the intersection of the original coalition polytope with two hyperplanes, i and j. The vector $(a_j \cdot v_1, ..., a_j \cdot v_{M_n - 1})$ is the normal vector to the $M_n - 2$ dimensional hyperplane formed by intersection of j-th and i-th hyperplanes in M_n dimensional space. By first order differentiability condition this vector is well defined.

For the second order derivative with respect to the same argument we have a slightly different formula. To derive it, observe that a change in b_i transforms into a change in x_0 in the set of inequalities 1.36. Above we defined x_0 as $a_i \cdot x_0 = b_i$. Then,

$$\frac{\partial (a_j \cdot x_0)}{\partial b_i} = \frac{a_i a_j}{||a_i||^2}, \tag{1.38}$$

which allows us to write down the second derivative as

$$\frac{\partial^2 V(M_n, A, b)}{\partial b_i \partial b_i} = \frac{1}{||a_i||} \sum_{j=1, j\neq i}^{I} \frac{-a_i a_j}{||a_i||^2} \frac{V_{ij}(M_n - 2, A, b)}{||(a_j \cdot v_1, ..., a_j \cdot v_{M_n - 1})||}. \tag{1.39}$$

Again, the continuity of derivative relies on the continuity of the volume function of a $M_n - 2$ dimensional polytope. Assuming that the continuity of first order derivatives is in place, we need to derive condition under which $M_n - 2$ dimensional polytope formed by intersection of $M_n - 1$ dimensional polytope with j-th hyperplane can be represented as intersection of halfspaces in $\mathbb{R}^{M_n - 2}$.

As before, consider a substitution

$$y = y_0 + \sum_{m=1}^{M_n - 2} z_m g_m + Z \frac{\nu_j}{||\nu_j||}, \tag{1.40}$$

where ν_j is a vector $\nu_j = (a_j \cdot v_1, ..., a_j \cdot v_{M_n - 1})$, vector y_0 is such that $\nu_j \cdot y_0 = b_j - a_j \cdot x_0$, and vectors g_m, for $m = 1, ..., M_n - 2$ form an orthonormal basis of subspace $\nu_j \cdot y = 0$ (which is also a subspace of $a_i \cdot x = 0$). Restating the set of inequalities 1.36 for $M_n - 1$ dimensional polytope,

$$\nu_j \cdot y \leq b_j - a_j \cdot x_0, \forall j \neq i, \tag{1.36}$$

and using substitution 1.40 we get

$$\begin{cases} 0 \leq -Z \frac{\nu j \cdot \nu_j}{||\nu_j||} \\ \sum_{m=1}^{M_n - 2} z_m (\nu_l \cdot g_m) \leq b_l - a_l \cdot x_0 - \nu_l \cdot y_0 - Z \frac{\nu_l \cdot \nu_j}{||\nu_j||}, \forall l \neq i, j. \end{cases} \tag{1.41}$$

For $Z = 0$ we get the domain of intersection of $M_n - 1$ dimensional polytope with j-th hyperplane.

$$\begin{cases} \sum_{m=1}^{M_n - 2} z_m (\nu_l \cdot g_m) \leq b_l - a_l \cdot x_0 - \nu_l \cdot y_0, \forall l \neq i, j \end{cases} \tag{1.42}$$

This domain is a $M_n - 2$ dimensional polytope in (A, b) representation with volume function continuous in b as long as $\forall l \ \exists m$ such that $\nu_l \cdot g_m \neq 0$. Thus, $V(M_n, A, b)$ is twice continuously differentiable under the conditions stated above.

□

Proof of proposition 4

The result is derived directly from equation 1.14 by dropping constants $\mathbb{V}(\theta_h; q_n^A, q_n^B)$ and $\bar{R}_i(\bar{q}_n^A)$; multiplying by δ and leaving out the error term. What remains is the linear approximation of part of candidate's A_n payoff that he has influence upon.

□

Proof of proposition 5

Consider the exact candidate A_n's objective function, see eq. 1.14. To establish concavity in policy choice variable, $q_n^{A_n}$, for fixed arbitrary values of all other inputs, I consider the declared above property of function $\bar{R}_i(q_n^{A_n})$,

$$\frac{\partial^2 R((q_1, ..., q_n, ..., q_N)^\intercal; \beta_i, \gamma_i)}{\partial q_n^2} < 0, \tag{1.43}$$

where the expectation is omitted as I look for a symmetric equilibrium. Then, the second derivative of $P[m_n = A_n](q_n^A | q_n^B)$ is

$$\begin{aligned} \frac{\partial^2 P[m_n = A_n](q_n^A | q_n^B)}{\partial (q_n^{A_n})^2} = \sum_{i=1}^{I} \frac{\bar{R}_i''(q_n^A)}{\delta} \sum_{h=1}^{\#\Theta_{A_n}} \frac{(2\theta_{hi} - 1)\mathbb{V}_i(\theta_h)}{||\sqrt{12} w^i \odot \sigma||} + \\ + \frac{\partial^2}{\partial (q_n^{A_n})^2} \max_i \left| \frac{\bar{R}_i(q_n^A) - \bar{R}_i(\bar{q}_n^A)}{\delta} \right|^2 O(1). \end{aligned} \tag{1.44}$$

Observe that

$$\sum_{h=1}^{\#\Theta_{A_n}} \frac{(2\theta_{hi} - 1)\mathbb{V}_i(\theta_h)}{||\sqrt{12} w^i \odot \sigma||} \geq 0$$

because for any winning coalition where investor-i is not present exists another coalition where investor-i is present. That coalition lies "right across the investor-i's hyperplane" so it exactly compensates the negative effect of the original one. Since $\mathbb{V}_i(\theta_h) \geq 0$ for all i, and because every manager has positive chance of being elected, I conclude that for at least one investor in at least one coalition, $\mathbb{V}_i(\theta_h) > 0$. That is the first term in equation 1.44 is negative. The second term in this equation has piecewise continuous second derivative with bounded absolute value. Since parameter δ enters as δ^2 in comparison to the first term, the impact of the second term can be made arbitrary small for large enough δ. That is, $\exists \underline{\delta}$ s.t. $\forall \delta > \underline{\delta}$ I have

$$\frac{\partial^2 P[m_n = A_n](q_n^A | q_n^B)}{\partial (q_n^{A_n})^2} < 0.$$

This implies concavity of objective function of a firm for large enough investors' taste shocks.

□

Proof of proposition 6

Recall, that for the equilibrium to exist, both investors and prospective managers need to play their best response strategies. I will tackle this problem sequentially in steps. First, I show that prospective managers reach an equilibrium with desired properties, given that investors behave in accordance to the results above. Second, I verify that investors' behavior is their best response to the prospective managers' actions.

Above, I derived the firm's objective function (equation 1.12) and established that under given policies of all other firms, prospective managers at an arbitrary firm n are going to play the same strategy. Moreover, I've shown that objective function is jointly continuous in policy parameters of all firms (see the proof of Proposition 3). Our next step is to show upper hemicontinuity of candidates' best responses.

For an arbitrary firm, n, consider a continuous mapping $\Gamma : [0, \bar{Q}]^{N-1} \to 2^{[0,\bar{Q}]}$ such that $\forall x \in [0, \bar{Q}]^{N-1}, \Gamma(x) = [0, \bar{Q}]$. Define function $\mathcal{M}(x) = \max_{q_n^{A_n}} \left\{ P[m_n = A_n](q_n^{A_n}) \middle| q_n^{A_n} \in \Gamma(x) \right\}$, where $q_{-n} = x$. Denote the best response correspondence as

$$\Phi(x) = \left\{ q_n^{A_n} \middle| q_n^{A_n} \in \Gamma(x), P[m_n = A_n](q_n^{A_n}) = \mathcal{M}(x) \right\}.$$

Then by applying Berge's Maximum theorem (Berge, 1963; Aliprantis et al., 2006) I establish that mapping $\Phi(x) : [0, \bar{Q}]^{N-1} \to 2^{[0,\bar{Q}]}$ is non-empty, upper hemi-continuous, and compact-valued. Moreover, since for sufficiently large δ firm's objective function is concave, the mapping $\Phi(x)$ is also convex-valued (see Sundaram (1996) p. 239).

Construct correspondence, $\mathcal{B} : [0, \bar{Q}]^N \to \times_{n=1}^{N} 2^{[0,\bar{Q}]}$, to combine best responses of prospective managers[24] at different firms as $\mathcal{B}(x) = \Phi_1(x) \times \Phi_2(x) \times ... \times \Phi_N(x)$, where $\times$ is Cartesian product. Here correspondences $\Phi_n(x)$ are the ones defined in the paragraph above, except that n-th correspondence ignores the firm-n's policy, hereby reducing the dimensionality of the argument to $N-1$. In Theorem 17.28 Aliprantis et al. (2006) show that a Cartesian product of finite family of upper hemicontinuous correspondences with compact values is upper hemicontinuous with compact values. Hence, correspondence $\mathcal{B}(x)$ is upper hemicontinuous and compact-valued. Since underlying correspondences are non-empty, $\mathcal{B}$ is non-empty as well. For high enough δ, correspondence $\mathcal{B}$ is convex-valued since Cartesian product preserves convexity.

Application of Kakutani's fixed-point theorem (Aliprantis et al., 2006) to correspondence $\mathcal{B}$ establishes existence of compact-valued non-empty set of fixed points,

$$\left\{ x \in [0, \bar{Q}]^N \middle| \mathcal{B}(x) = x \right\} \neq \varnothing.$$

Now, what I have shown is the existence of a fixed point for prospective managers' best responses to their own and investors' actions. We need to demonstrate that investors are best responding to managers and themselves as well. Given that the fixed point features symmetric candidates' proposals within any firm, using assumption 1, I observe that lemma 3

[24] Here I use the result that if equilibrium exists, prospective managers of a firm play the same strategies.

gives rise to inequality 1.5 as shown in chapter 1.3. This inequality represents best response at firm-n of an investor-i to prospective managers' equilibrium strategies at all firms expect firm-n where candidates A_n and B_n may play deviated strategies. Lemma 3 establishes independence of investor-i's best response from other investors' actions. This effectively implies *sincerity* of investors' best responses. Then, geometric analysis leads to the set of inequalities 1.6, which describes the probability of support coming from a single coalition. Finally, taking a sum over all wining coalitions, Θ_{A_n}, I receive equation 1.10 that describes prospective manager A_n's objective function. This completes the loop, and thus I have shown the existence of a Subgame Perfect Nash equilibrium.

□

Chapter 2

A Connection Between Shareholders' Portfolios and Directors' Tenures

2.1 Introduction

Do corporate elections matter? Every year millions of shareholders participate in tens of thousands of elections at thousands of shareholder meetings in the U.S. Yet less than 1% of director elections are contested,[1] more than 99% of auditor ratification votes receive support from greater than a half of voting shares outstanding,[2] and an ever rising block of shares is voted by widely diversified institutional investors.[3] In this paper, I analyze the structure of shareholder meetings, the effect of shareholder support on directors' tenure, and whether mutual funds take their portfolios into consideration when making voting decisions. I find that the level of shareholder support is positively associated with the length of director's tenure at a company and the probability of future nomination. I examine whether fund mergers, a possible source of shocks to mutual funds' portfolio composition, affect mutual

[1] Institutional Shareholder Services Voting Analytics database documents more than 18000 director election events in 2015 and only 107 of those correspond to directors in opposition slates.

[2] Data comes from ISS Voting Analytics database.

[3] Backus et al. (2019b)

funds' voting patterns. In an event study, I observe that a merger affects the acquiring mutual fund's voting behavior.

First, I study the composition of election issues at shareholder meetings. In a frequency analysis, I find that director elections and auditor ratifications make up more than 80% of all election issues and appear in 94% and 80% of all shareholder meetings respectively. A meeting's agenda composition and other meeting characteristics, like location and time, might affect shareholder participation (Van der Elst, 2011; Li & Yermack, 2016). Since a loss of shareholder support may partially come from non-participation of some shareholders (Nili & Kastiel, 2016; Jill E. Fisch, 2017; Cvijanovic et al., 2019), I choose director elections, as the most uniformly present election type, to study the effect of shareholder approval on. In an application of principal component analysis to the classification of shareholder meetings' compositions, I obtain a similar result: the presence of director elections on meeting agenda is one of the least important factors in distinguishing between different composition types of those meetings. This finding strengthens my view that director elections face one of the most representative samples of voting shareholders.

Second, I conduct a short review of competitiveness for the first few of the most popular election issue types. Using the data on director elections, I observe that in less than 1% of cases, a management-proposed slate of director nominees is contested by an opposition slate. Moreover, Cai et al. (2013) find that the rare instances of director loosing an election almost always result in him/her staying on the board. A similar situation holds for auditor ratifications: in a less than 1% of ratification votes an auditor fails to secure more than a half of the outstanding voting shares as votes of support. Less numerous compensation election issues demonstrate a greater disagreement in shareholders' votes: in 12% of cases "Say-On-Pay" proposals receive support from less than a half of all voting shares outstanding.

Third, I focus on uncontested director elections as the most uniformly present agenda issue at shareholder meetings. Literature suggests that bad election performance at uncontested director elections does not lead to an immediate removal of the director from the board (Cai et al., 2009, 2013), yet there is evidence of a delayed effect (Iliev et al., 2015; Aggarwal et al., 2019). I test two hypotheses: *(a)* directors with low shareholder support refrain or are prevented from being nominated in the next election cycle at the company; and *(b)* directors with low shareholder support experience shorter employment spells at the company. I find both hypotheses to be supported by the data.

For the former hypothesis, using a logit regression, I find that a higher fraction of votes "For" out of all voting shares outstanding has a positive and significant relationship with the probability of director's nomination in the future. Aggarwal et al. (2019) find that greater percentage of dissent votes is related to higher chances of departing from the board, which is consistent with my results. My approach differs from theirs in capturing the effect of both dissent votes and shareholder apathy.[4]

I test the latter hypothesis using a time-varying Cox's proportional hazard model. I find that low shareholder support is associated with shorter director's tenure at the company. Moreover, the magnitude of the effect is amplified by majority voting requirement, board being staggered, and a positive ISS recommendation.

Fourth, I study the effect of mutual funds mergers on an acquiring funds' voting behavior. I find that a fund alters its voting pattern soon after a merger with another fund. Because mergers reshape the acquiring fund's portfolio structure, this study suggests that there might be an effect of portfolio structure on the voting behavior. Yet, mergers may also involve other adjustments to the acquiring fund that might cause this change in its voting behavior.

I compare a fund's voting record with voting records of other funds at the same firms. This approach presents two hurdles: *(i)* since mutual funds are very diverse in terms of their portfolios and voting behaviors, a direct comparison with any single mutual fund is almost never possible; and *(ii)* a comparison based on a group of mutual funds typically leads to a missing data problem due to variation in the funds' portfolios. I overcome the first hurdle by constructing an artificial mutual fund, a synthetic control (Abadie et al., 2010), to which

[4] I attribute shareholders' apathy to the non-voted shares.

I then compare the voting behavior of the actual fund. To my knowledge, I'm the first to use the synthetic control method in studying corporate governance.[5] To deal with the second hurdle, I use the robust version of the synthetic control method proposed by Amjad et al. (2018).

Focusing on voting behavior of acquiring mutual funds before and after a merger, I observe a significant change in their voting patterns. To alleviate concerns that the effect might be caused by the timing of the mergers before/after the bulk of corporate elections happen or that it is just an artefact of the synthetic control method, I conduct a placebo study where I use the same set of merger dates but I replace the acquiring mutual funds with arbitrary non-merging ones. The placebo study shows no change in voting behavior which rules out the aforementioned concerns.

The dependence of voting behavior on portfolio structure is important because it contradicts the Fisher separation theorem (Fisher, 1930). The theorem establishes that shareholders should unanimously agree on the firm's profit maximizing production plan. The evidence of it not happening suggests that the price-taking assumption is likely violated and that mutual funds are internalizing the externalities that firms place on each other.

2.2 Election types

Shareholder meetings include a wide range of election agendas, yet more than 90% of all election issues fall into one of the top ten popular types. Director elections is by far the most popular election type: 7 out of 10 election issues are in this category. Typically, there are multiple directors up for election at a given annual shareholder meeting which inflates the number of director elections in comparison to other election issues. Table 2.1 presents counts of the most common election types.

[5] The synthetic control method was pioneered by Abadie et al. (2010) in their study of California's tobacco control program. The method gone largely unnoticed in finance literature (Atanasov & Black, 2016). A notable exception is the paper by Berger et al. (2020). Recently, the synthetic control method gained traction in voting behavior studies.

Table 2.1: Election agendas at shareholder meetings. Data comes from ISS Voting Analytics dataset for years 2003 - 2016.

Agenda general description	Count	Cumulative %
Elect Director	234305	69.64
Ratify Auditors	29577	78.43
Advisory Vote to Ratify Named Executive Office...	15323	82.99
Amend Omnibus Stock Plan	7025	85.08
Ratify X as Auditors	5669	86.76
Approve Omnibus Stock Plan	4122	87.99
Advisory Vote on Say on Pay Frequency	3536	89.04
Approve/Amend Executive Incentive Bonus Plan	2811	89.87
Elect Subsidiary Director	2465	90.61
Increase Authorized Common Stock	1885	91.17
Adjourn Meeting	1774	91.69
Amend Qualified Employee Stock Purchase Plan	1408	92.11
Declassify the Board of Directors	1360	92.52
Approve Merger Agreement	1330	92.91
Other Business	1072	93.23
Elect Director (Management)	930	93.51
Elect Directors (Opposition Slate)	903	93.77
Amend Articles/Bylaws/Charter-Non-Routine	882	94.04
Approve Qualified Employee Stock Purchase Plan	756	94.26
Reduce Supermajority Vote Requirement	643	94.45
Amend Stock Option Plan	624	94.64
Require Independent Board Chairman	608	94.82
Advisory Vote on Golden Parachutes	514	94.97
Political Contributions Disclosure	508	95.12
Require a Majority Vote for the Election of Di...	477	95.26
Approve Acquisition OR Issue Shares in Connect...	463	95.40
Approve Auditors and Authorize Board to Fix Th...	441	95.53
Amend Non-Employee Director Stock Option Plan	357	95.64
Company Specific-Equity-Related	336	95.74
Approve Reverse Stock Split	292	95.83
Change Company Name	282	95.91
Amend Articles/Bylaws/Charter – Call Special ...	274	95.99
Approve Stock Option Plan	266	96.07
Amend Non-Employee Director Omnibus Stock Plan	265	96.15
Company-Specific – Shareholder Miscellaneous	255	96.22
Restore or Provide for Cumulative Voting	236	96.29
Provide Right to Act by Written Consent	225	96.36
Approve Non-Employee Director Omnibus Stock Plan	223	96.43
Approve Repricing of Options	222	96.49
Proxy Access	219	96.56

While director elections comprise an overwhelming majority of issues recorded, other categories also appear frequently at shareholder meetings. In table 2.2, I present 40 most frequent election types at these meetings. Director elections appear in 94% of the meetings, followed closely by ratifying auditors at 80% of the meetings. Other frequent issues are compensation questions,[6] governance questions,[7] merger issues,[8] and proposals related to equity.[9]

2.2.1 Election types correlation

Companies follow a certain agenda structure at shareholder meetings. To discover that structure, I first study the correlation between different types of popular agenda items, typically present at shareholder meetings. Second, I use principal component analysis to characterize what components of shareholder meetings define the most important dimensions across which the content of one shareholder meeting is different from another.

Composition of shareholder meetings is studied in figure 2.1. Since companies may use somewhat different titles for similar agenda items, I classify all items that appear in at least 0.5% of meetings into 7 major purpose groups.[10] Then, I study how these groups correlate in appearance at shareholder meetings.

[6] These include issues such as "Advisory Vote to Ratify Named Executive Officers' Compensation", "Amend Omnibus Stock Plan", "Approve Omnibus Stock Plan", "Advisory Vote on Say on Pay Frequency", "Approve/Amend Executive Incentive Bonus Plan", "Amend Qualified Employee Stock Purchase Plan", "Approve Qualified Employee Stock Purchase Plan", "Amend Stock Option Plan", "Advisory Vote on Golden Parachutes", "Amend Non-Employee Director Stock Option Plan", "Amend Non-Employee Director Omnibus Stock Plan", "Approve Stock Option Plan", "Approve Non-Employee Director Omnibus Stock Plan", "Approve Repricing of Options", "Performance-Based and/or Time-Based Equity Awards", among others.

[7] These contain such agendas as "Declassify the Board of Directors", "Require Independent Board Chairman", "Political Contributions Disclosure", "Amend Articles/Bylaws/Charter-Non-Routine", "Require a Majority Vote for the Election of Directors", "Reduce Supermajority Vote Requirement", "Amend Articles/Bylaws/Charter – Call Special Meetings", "Restore or Provide for Cumulative Voting", "Proxy Access", among others.

[8] Merger issues have "Approve Merger Agreement" and "Approve Acquisition OR Issue Shares in Connection with Acquisition" agenda names.

[9] Most common agenda names are "Stock Retention/Holding Period", "Increase Authorized Common Stock", "Company Specific-Equity-Related", and "Approve Reverse Stock Split".

[10] The exact composition of these groups is provided in the beginning of this chapter.

Table 2.2: Frequencies of occurrence of 40 most popular election agendas at shareholder meetings. Data comes from ISS Voting Analytics dataset for years 2003 - 2016.

Frequency	Agenda
0.9351	Elect Director
0.6751	Ratify Auditors
0.3493	Advisory Vote to Ratify Named Executive Office...
0.1556	Amend Omnibus Stock Plan
0.1287	Ratify X as Auditors
0.0937	Approve Omnibus Stock Plan
0.0810	Advisory Vote on Say on Pay Frequency
0.0619	Approve/Amend Executive Incentive Bonus Plan
0.0430	Increase Authorized Common Stock
0.0402	Adjourn Meeting
0.0319	Amend Qualified Employee Stock Purchase Plan
0.0305	Declassify the Board of Directors
0.0302	Approve Merger Agreement
0.0246	Other Business
0.0173	Approve Qualified Employee Stock Purchase Plan
0.0140	Amend Articles/Bylaws/Charter-Non-Routine
0.0139	Require Independent Board Chairman
0.0135	Amend Stock Option Plan
0.0116	Advisory Vote on Golden Parachutes
0.0112	Political Contributions Disclosure
0.0108	Require a Majority Vote for the Election of Di...
0.0103	Approve Acquisition OR Issue Shares in Connect...
0.0098	Approve Auditors and Authorize Board to Fix Th...
0.0080	Amend Non-Employee Director Stock Option Plan
0.0080	Reduce Supermajority Vote Requirement
0.0068	Company Specific-Equity-Related
0.0064	Change Company Name
0.0062	Amend Articles/Bylaws/Charter – Call Special ...
0.0061	Approve Reverse Stock Split
0.0060	Amend Non-Employee Director Omnibus Stock Plan
0.0058	Approve Stock Option Plan
0.0054	Restore or Provide for Cumulative Voting
0.0051	Provide Right to Act by Written Consent
0.0051	Approve Non-Employee Director Omnibus Stock Plan
0.0050	Proxy Access
0.0050	Company-Specific – Shareholder Miscellaneous
0.0048	Approve Repricing of Options
0.0046	Stock Retention/Holding Period
0.0046	Submit Shareholder Rights Plan (Poison Pill) t...
0.0042	Performance-Based and/or Time-Based Equity Awards

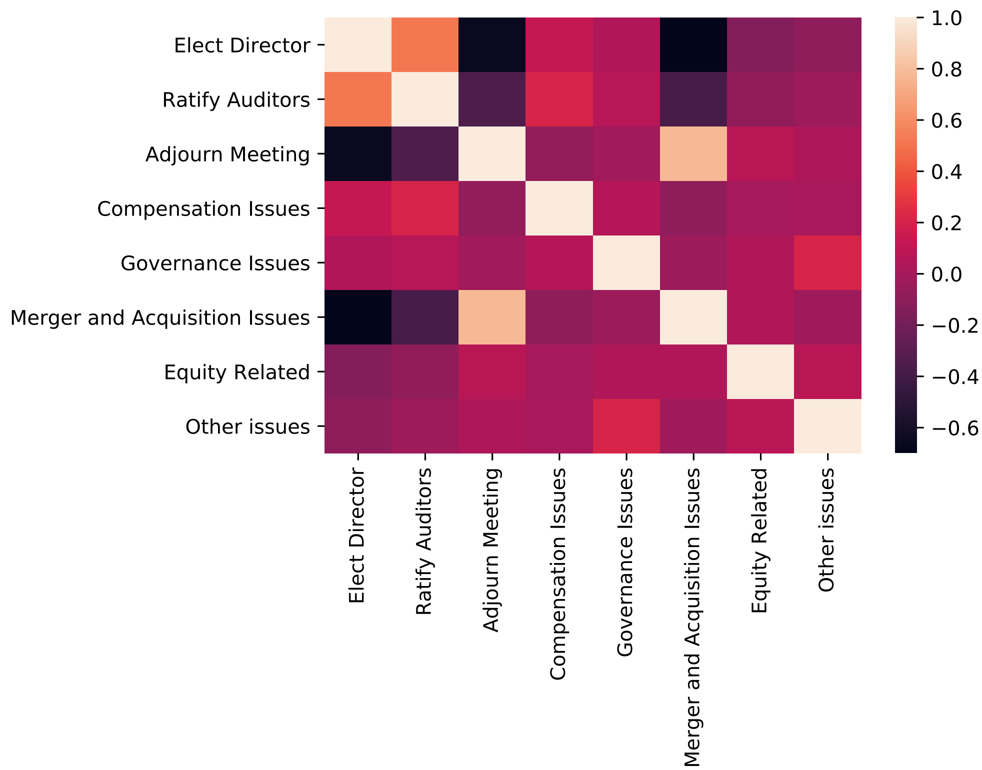

Figure 2.1: Correlation among the most frequent election agendas at shareholders' meetings.

Correlation structure reveals two major kinds of shareholder meetings. The first kind includes annual meetings where shareholders elect directors and ratify auditors. This type is also likely to include questions on compensation and corporate governance. The second kind includes special meetings (non-regular, proxy contest, etc.) that deal with merger and acquisition issues as well as contested director elections. The majority of shareholder meetings falls into the first category, with a much smaller portion falling into the second category. Some shareholder meetings exhibit features of both categories and cannot be easily classified into one specific kind.

2.2.2 Principal component analysis of shareholder meetings

Principal component analysis (PCA) allows me to mechanically discover the combinations of elections issues that differ the most between different kinds of shareholder meetings. In this section, I apply PCA to identify those combinations (principal components), to perform cluster analysis to identify distinct kinds of shareholder meetings, and to describe the differences between these kinds using principal component loadings. I find that shareholder meetings can be classified into four major categories: meetings which contain compensation and infrequent[11] election issues, meetings that contain none of those, and meetings that contain one or the other type of aforementioned election issues.

To perform PCA for shareholder meetings composition I implement the following steps. First, for each shareholder meeting, I construct a vector that describes the types of elections present at the meeting. A component of this vector is equal to 1 if one or more elections of the corresponding type were conducted at the meeting and is equal to 0 if otherwise. The length of the vector is equal to the number of election types considered.[12] Second, for a collection of such vectors for all shareholder meetings, I perform mean removal for individual components. Since, by construction, each component has the same scale, I do not normalize the components and leave the variance of individual components unchanged. This approach enhances PCA's search for directions of highest variance by accounting for the natural differences in variances of individual vector's components.

Principal component analysis reveals that shareholder meetings can be classified into four major categories. While PCA does not endow those categories with a meaningful explanation of differences between them, a certain level of understanding can be reached by studying eigenvectors' (principal components) loadings. Using a graphical representation of PCA results for the first three principal components (see a score plot in fig. 2.2), I observe that four shareholder meeting categories arise from two splits. First, shareholder meetings are divided into two categories by the presence of compensation issues. Second, those two categories are subdivided in two halves each by the presence of infrequent election issues

[11] Infrequent election agendas typically include shareholder-proposed and firm-specific issues.

[12] Here, I consider the same election types that were used in the previous section in construction of the correlation diagram. While it is possible to perform PCA on the whole set of 389 different election types' descriptions available from the ISS Voting Analytics database, this would skew the results as descriptions are not perfect. The dataset contains multiple mutually excluding election descriptions (e.g. "Ratify Auditors" and "Ratify X as Auditors") that would cause artificial clustering of shareholder meetings as only one such description is used at a given shareholder meeting.

(these are typically firm-specific and shareholder-proposed issues). Very frequent election issues, such as director elections and auditor ratifications, do not play a substantial role in differentiating between various categories of shareholder meetings.

The first split happens in the space spanned by the two major principal components, and the second split happens along the third most significant principal component. The top left chart in figure 2.2 shows the first split. Two point clouds represent shareholder meetings classified into different categories by the PCA. A notable feature of this split is that clouds' internal structure is very similar. This could be explained in the following way: while categories split is driven by a major factor that is aligned well with the subspace spanned by first two principal components, other factors that explain less pronounced differences between shareholder meetings are not aligned well with this subspace. Therefore, if these minor factors are independent of the major factor that drives the split, then both point clouds should have similar structures. Clouds' structures are determined by the projection of the minor factors on the subspace defined by the major principal components. Thus, I can identify the direction of the major factor's influence within the principal components' subspace as a vector, that if added to the one cloud's points would shift this cloud to overlap with the other one.

Using coherent point drift algorithm by Myronenko & Song (2010), I match the point clouds from the upper left plot in figure 2.2 and find that the difference between categories in the first split corresponds to a shift of the red cloud along the vector $[-0.86, 0.49]$ in the space of the first two principal components. Table 2.3 presents loadings of principal components on original variables. Using these loadings and the vector $[-0.86, 0.49]$, I compute loadings of the difference between categories as weighted combination of the first two principal components. The result is presented in table 2.4. A substantially higher weight is placed on compensation issues dummy. This result signals that inclusion of compensation issue(s) on shareholder meeting's agenda list produces the highest variation in shareholder meeting composition.

Table 2.3: Principal components' loadings for the Principal Component Analysis (PCA) of shareholder meetings' compositions. Original variables are sorted in a way that shows their importance according to the PCA: for each eigenvector I determine the most important original variable based on absolute loadings, then I sort the original variables by the order of eigenvectors based on this relation.

	Eigenvectors							
Original dummy variable	1	2	3	4	5	6	7	8
Compensation Issues	-0.86	0.49	-0.16	-0.03	-0.05	-0.01	-0.01	0.00
Ratify Auditors	-0.44	-0.64	0.22	0.07	0.58	-0.01	-0.10	-0.01
Other issues	0.01	0.31	0.90	-0.29	0.04	-0.06	0.04	0.02
Governance Issues	-0.05	0.05	0.29	0.93	-0.21	-0.06	-0.01	0.00
Equity Related	0.02	0.09	0.04	0.07	0.12	0.98	0.06	0.03
Elect Director	-0.20	-0.36	0.08	-0.10	-0.43	0.04	0.78	0.13
Merger and Acquisition Issues	0.13	0.25	-0.11	0.13	0.45	-0.13	0.30	0.76
Adjourn Meeting	0.12	0.25	-0.08	0.12	0.46	-0.10	0.53	-0.63

Table 2.4: Loadings of the difference vector between shareholder meeting categories in the first split produced by the PCA. These loadings are computed as a weighted sum of the first two major principal components' loadings with vector $[-0.86, 0.49]$ providing the weights. Vector $[-0.86, 0.49]$ represents a difference between shareholder meeting categories in the subspace spanned by the first two principal components. The vector specifies a shift of red point cloud needed such that it would cover blue point cloud in the upper left plot in figure 2.2.

Original dummy variable	Weight	Absolute Weight
Compensation Issues	0.9751	0.9751
Other issues	0.1377	0.1377
Governance Issues	0.0686	0.0686
Ratify Auditors	0.0639	0.0639
Equity Related	0.0247	0.0247
Adjourn Meeting	0.0164	0.0164
Merger and Acquisition Issues	0.0102	0.0102
Elect Director	-0.0078	0.0078

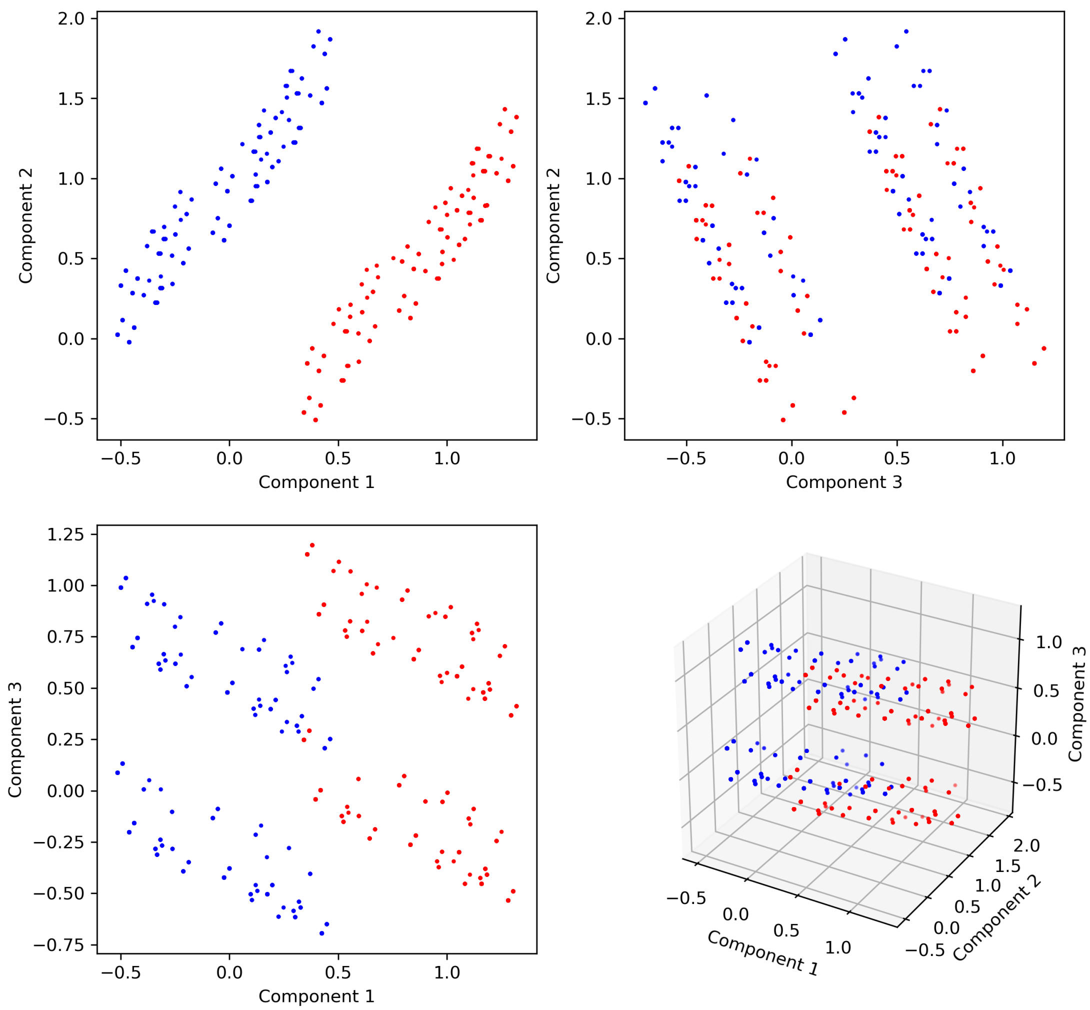

Figure 2.2: Score plot for the principle component analysis of shareholder meetings' composition. Components are ordered by the share of variance explained in descending order. The top left chart shows shareholder meetings' scores for the first two principle components. Two clusters are clearly present. Clusters are color-coded differently to track their evolution across projections at other combinations of principal components. Top right and bottom left charts present projections involving the third principal component. Both charts show that the third principal component splits the original clusters. Therefore, 4 different clusters are identifiable in the space spanned by the first three principal components.

Table 2.5: Summary statistics for shareholder meetings' agenda items. Variables are equal to 1 if corresponding agenda items were present at a shareholder meeting and 0 otherwise.

Agenda item present	mean	std
Elect Director	0.9387	0.2398
Ratify Auditors	0.8134	0.3896
Compensation Issues	0.5815	0.4933
Other issues	0.1780	0.3825
Governance Issues	0.0861	0.2805
Equity Related	0.0585	0.2346
Adjourn Meeting	0.0402	0.1965
Merger and Acquisition Issues	0.0401	0.1962

Second split happens mostly along the dimension of the third principal component. This split is driven by the less frequent election agenda issues that are not classified in the short list I'm using in this paper. While such issues individually have a frequency of occurrence at shareholder meetings of less than 0.5%, together those issues appear in 17.8% of meetings. This substantial proportion results in high standard deviation in presence of those rare election issues.[13] Therefore, presence of those election issues serves as a good differentiating factor between shareholder meeting categories.

Election issues with high variance of occurrence are more likely to be distinctive features of shareholder meetings. Table 2.5 presents summary statistics for shareholder meetings' agenda item types. Both compensation and "other" issues have comparatively high variance. Very common issues, like director elections, and quite rare, like equity related and merger and acquisition issues, have relatively small variances. Those are less likely to drive the separation of shareholder meetings into different categories. I do not conduct graphical analysis of shareholder meetings clustering beyond the first three principal components. Yet, principal component loadings in table 2.3 shed some light on influence of other election agenda types on meetings' composition.

[13]Since all original variables here are dummies, the standard deviation is directly related to the mean as per Bernoulli distribution: $\sigma = \sqrt{p(1-p)}$. Standard deviation would be highest for election issues that were present in exactly half of all shareholder meetings.

Governance issues comprise a significant portion of the fourth principal component's vector. Having smaller standard deviation than compensation or infrequent issues, governance issues play a modest role in differentiation of shareholder meetings. Auditor ratifications and equity related issues are predominantly aligned with fifth and sixth principal components vectors respectively. While auditor ratifications also appear in the first two principal component's vectors, analysis above clearly shows that they do not play a significant role in the first two splits between shareholder meetings' categories. Scree plot in figure 2.3 shows that fifth and sixth principal components explain proportion of original data's variance similar to the fourth principal component. Therefore, auditor ratifications, equity, and governance related issues are equally important in composition of shareholder meeting.

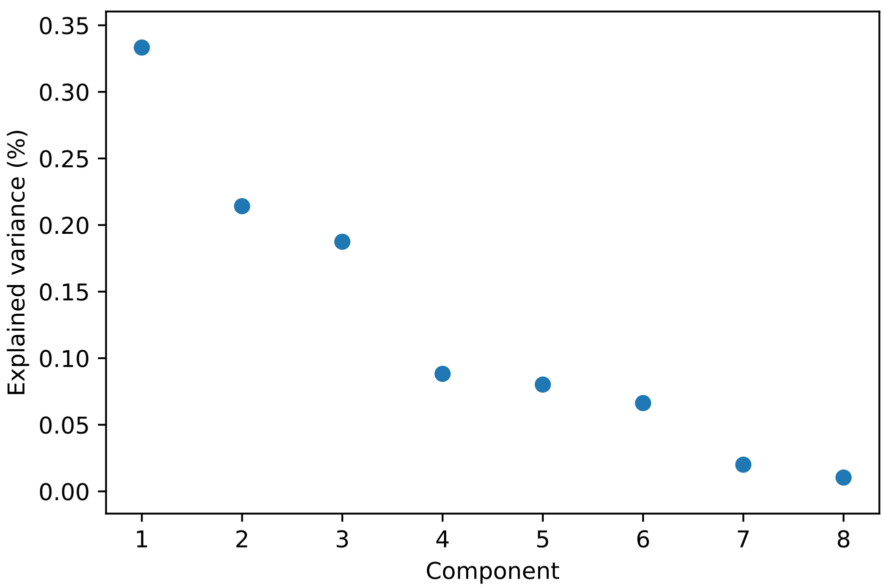

Figure 2.3: Scree plot for principle component analysis of shareholder meetings' composition. The first three components explain more than 73% of variance in the original data.

Lastly, director elections and merger and acquisition issues are mostly pronounced in the last two principal components' vectors. These issues do not introduce much differentiation in shareholder meetings. Director elections are present in 94% of shareholder meetings, while mergers and acquisitions get voted on in only 4% of meetings. Therefore, both types of elections have relatively small variance of occurrence, and they explain only a limited amount of variance in shareholder meetings' composition. Mergers and acquisition issues typically get accompanied by adjourn meeting votes. Figure 2.1 shows substantial positive correlation between these two types of elections issues. Adjourn meeting votes do not represent the biggest element in any one of the principal components' vectors. Thus it is likely that they do not contribute much to differentiation in shareholder meetings' composition either.

2.3 Electoral competitiveness

Corporate elections are not always binding, few are contested, and shareholder proposals are infrequent. Most popular agenda items, like director elections and auditor ratifications, do not provide a significant variation in outcomes. In fewer than 5% of cases for director elections and 1% for auditor ratifications proposals fail to secure votes in favor from more than a half of all outstanding voting shares. Less frequent issues, like "Say-On-Pay" votes, tend to feature greater share of cases with low shareholder support.

2.3.1 Director elections

Director elections are the most common type of election issues at shareholder meetings. Approximately 70% of all elections are of this type and about 94% of all shareholder meetings involve director elections. Yet, competitiveness of director elections is rather low: overwhelming majority of all director elections are uncontested, a significant portion of companies use plurality voting standard,[14] and very few director nominees lose elections and even then most of them become directors.

Director elections typically happen at annual shareholder meetings where a slate of director nominees is proposed for an election. In a very small number of cases, less than 0.4%, two slates are proposed: by the management and by the opposition. Companies use two methods

[14]Under a plurality voting standard, an uncontested nominee needs just one vote in favor in order to be elected.

for electing directors: majority and plurality voting standards. Under the plurality voting standard, a director nominee receiving the most votes in support wins the election. This method has a substantial drawback: in an uncontested election the nominee needs to receive just one vote "For" in order to get elected. An alternative method is the majority voting standard. This approach requires a winning nominee to secure more votes in support than votes against. Figure 2.4 shows the share of director elections requiring a winning nominee to pass 50% (or above) threshold. The growth in the last decade is attributed to the changes in voting practices of the S&P 500 companies, while smaller companies from Russell 2000 mostly stick to the plurality voting standard (see Council of Institutional Investors (2017) for details).

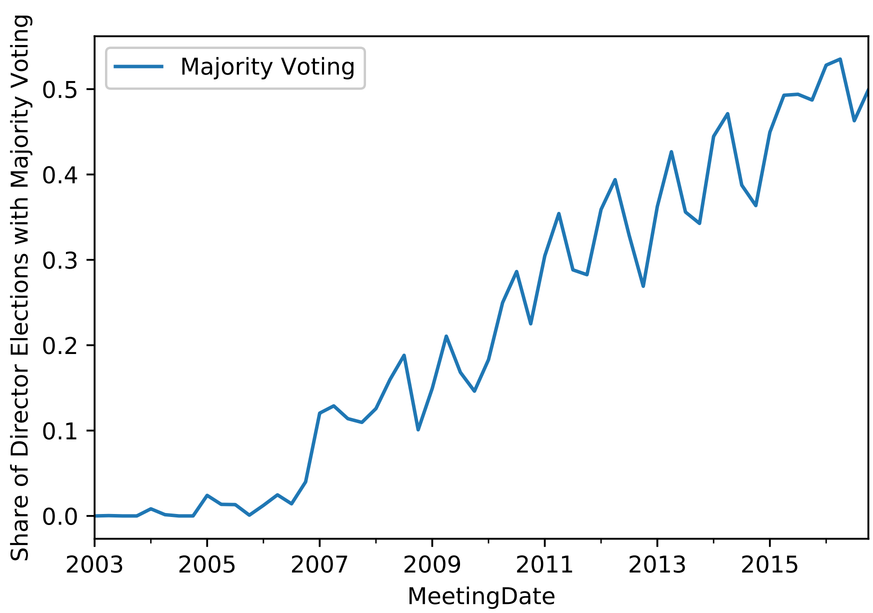

Figure 2.4: Share of director elections requiring winning nominee to pass 50% (or above) threshold. Data on threshold comes from the ISS Voting Analytics dataset. I compute the share by dividing the number of director elections with the threshold requirement by the total number of director elections in a given quarter.

Director nominees almost always enjoy a high support rate. I define support rate as the proportion of shares voted "For" to the total number of voting shares outstanding. This

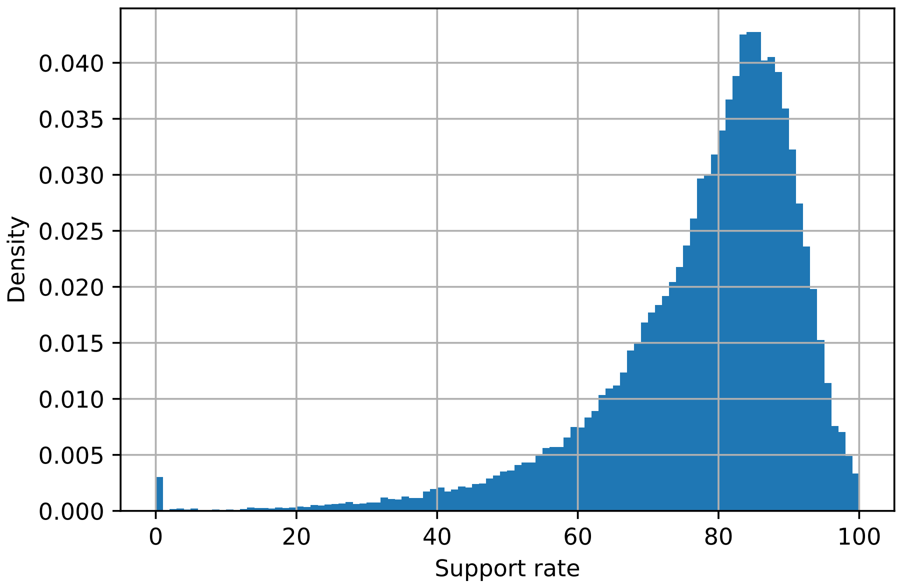

Figure 2.5: Density histogram of support rates at director elections. Support rate is defined as the ratio of votes cast "For" to the total number of shares outstanding (voting shares). Only 4.7% of director elections had support rate of less than 50% of shares outstanding.

is a very conservative estimate, as less than 1% of companies use total shares outstanding as the base in their calculations. Figure 2.5 presents a density histogram for support rates at director elections. The distribution has a heavy rightward skew. Only 4.7% of director elections received support of less than half of outstanding shares. Since most companies in the sample do not use the majority voting standard and the ones that do typically use votes cast as the base in their calculations, the chances of a director nominee to miss the threshold are even smaller. Cai et al. (2013) investigate director elections from 2004 to 2010 and find that from 105445 directors only 294 directors at 153 firms received less than 50% of votes "For". Moreover, only 14 firms adopted some form of the majority voting standard. Therefore, even if a company uses the majority voting standard, there is a very slim chance that a nominee does not meet the threshold.

2.3.2 Auditor ratifications

The second most frequent corporate elections' issue is seeking ratification from shareholders on auditor selection. Over 80% of shareholder meetings in the dataset include auditor ratifications. This election issue is widespread partially due to the Sarbanes–Oxley Act of 2002 that introduced new auditor approval requirements. The act requires exchange-listed companies to have an audit committee that appoints and oversees the auditor. Corresponding SEC rules (SEC, 2003b) permit shareholder ratification of auditor selection in compliance with Sarbanes–Oxley Act. The idea behind the ratification process is that shareholders could voice their concern with the audit committee's work, selection of a specific auditor, and audit fees. Auditor ratifications are non-binding.

The density distribution of the support rate[15] for auditor ratifications has a significant rightward skew. Less than 1% of ratification elections do not receive at least half of voting shares outstanding in "For" votes. Therefore, virtually all auditors put up for a ratification are successfully ratified by shareholders.

In 2009 SEC allowed NYSE to update Rule 452 (SEC, 2009). The update prohibited brokers from voting clients' shares in uncontested director elections if they did not receive voting instructions from their clients. Unlike with director elections, brokers were not prohibited to vote shares in auditor ratifications (and other routine matters) without clients' instructions. This might have incentivized firms to include auditor ratifications on shareholder meetings in order to reach a quorum.

Krishnan & Ye (2005) find that companies might avoid auditor ratification when shareholders are dissatisfied with the boards of directors. They also point out that the likelihood of including an auditor ratification to a shareholder meeting is positively associated with financial expertise of audit committees. Dao et al. (2008) show that shareholders are more

[15] As before, I define the support rate as proportion of "For" votes among the the total voting shares outstanding.

likely to withdraw their support at an auditor ratification election when the auditor's tenure at the company is long.

While there is little evidence that auditors formally fail ratification elections at companies, low support rates and high levels of shareholder dissent may affect auditors' dismissals and resignations. Sainty et al. (2002), using pre Sarbanes–Oxley Act data, provide evidence that high degree of investor dissatisfaction is associated with a firm being more likely to change auditors. Barua et al. (2017) examine auditor dismissals using the auditor ratification voting data from 2011 to 2014. They find that the proportion of shareholder votes against auditors ratifications is associated with subsequent auditor dismissals.

2.3.3 Compensation issues

Votes of shareholder approval of executive compensation, "Say-On-Pay" votes, comprise the third largest category among corporate election issues. In 2011, the SEC introduced changes to Section 14A of the Securities Exchange Act of 1934 that require public companies to hold an advisory vote on compensation of company's named executive officers (NEO). The change was mandated by the Dodd-Frank Act, a comprehensive reform of financial regulation in the U.S. This change led to a tenfold increase in the number of "Say-On-Pay" votes conducted at shareholder meetings: for instance, the sample contains 215 such votes in 2010 and 2829 in 2011. This vote is meant to be an annual check of executive officers' compensation and is non-binding.

The practical purpose of this vote is to allow shareholders to voice their concerns with the level of NEO's compensation, NEO's performance during the past year, and to convey that information directly to firm's management. Literature finds that "Say-On-Pay" votes are beneficial for the firm. Iliev & Vitanova (2019) show that regular "Say-On-Pay" votes are valuable to shareholders. Cuñat et al. (2016) use pre Dodd-Frank era data to find that adoption of "Say-On-Pay" proposal leads to a 5% increase in the market value of a company. Robin Ferracone & Dayna Harris (2011) provide evidence that pay for performance disconnect, poor pay practices, and poor disclosure were the most common reasons to vote against in the failed "Say-On-Pay" votes in post Dodd-Frank era. Cotter et al. (2013) find that companies with low total shareholder return, inadequately high levels of executive pay, and companies with negative recommendations from ISS were faced with greater shareholder dissent at "Say-On-Pay" elections. The authors also note that despite the non-binding nature of these elections, companies that failed a vote undertook a change in their compensation schedules or engaged in additional communication with shareholders.

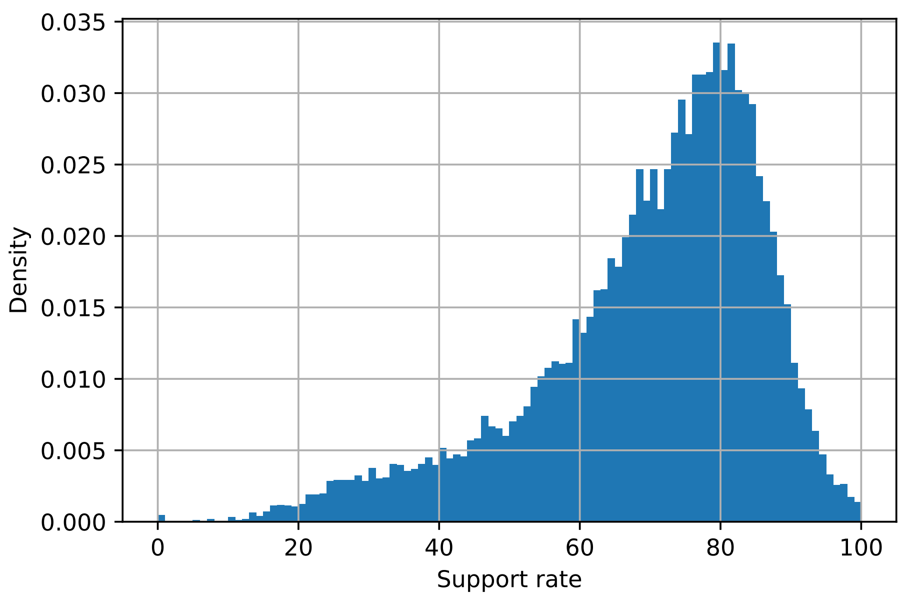

Figure 2.6: Density histogram of support rates at "Say-On-Pay" elections. Support rate is defined as the ratio of votes cast "For" to the total number of shares outstanding (voting shares). More than 12% of "Say-On-Pay" elections had support rate of less than 50% of shares outstanding.

Figure 2.6 presents density histogram of "Say-On-Pay" votes' outcomes. Unlike director elections and auditor ratifications, "Say-On-Pay" elections have a sizeable share of cases (12%) where proposal was not supported by a half of all voting shares outstanding.

2.4 Election results and director's job security

Director nominees very often receive high support rates in uncontested elections which makes losing an election highly improbable even under the majority voting standard. This might create an impression that uncontested director elections are a pure formality that does not affect nominee's chances of being elected as a director (Monks & Minow, 2004). In this section, I evaluate how election results are related to elected director's career prospects with a company. I find that low shareholder support more often precedes director's departure from the company. In a time-invariant Cox's proportional hazard model I do not find a significant relationship between average level of shareholder support and the length of director's tenure. When I account for the evolution of shareholder support rate and other covariates in a time-varying Cox's model, I find a significant and economically meaningful association between the support rate and the length of director's employment.

Directors value their reputation and might react in the loom of a low support vote (Grundfest, 1993). For contested elections, literature provides evidence that contested directors face a reduction in the number of directorships both in the targeted company and non-targeted companies (Fos & Tsoutsoura, 2014). Due to a high cost for activist shareholder, contested elections comprise less than 1% of all director elections. While uncontested elections do not pose a credible direct threat to nominees of losing directorships, elections' results serve as a signal of shareholder perception of the board and CEO performance. Fischer et al. (2009) find that firms with low board approval rates are associated with greater board and CEO turnover and lower CEO compensation. Guercio et al. (2008) study "just vote no" campaigns and find that such concerted actions of shareholders motivate boards to act in shareholders' interests.

2.4.1 Directorships' spells and election results data

Directors typically stay with a company for a number of years and participate in multiple elections. For the best possible coverage of election events, I reconstruct employment spells' lengths from the election data. I find that about a third of directors serve on staggered boards, while others participate in elections annually. Some directors experience transition of their boards between staggered/non-staggered structures.

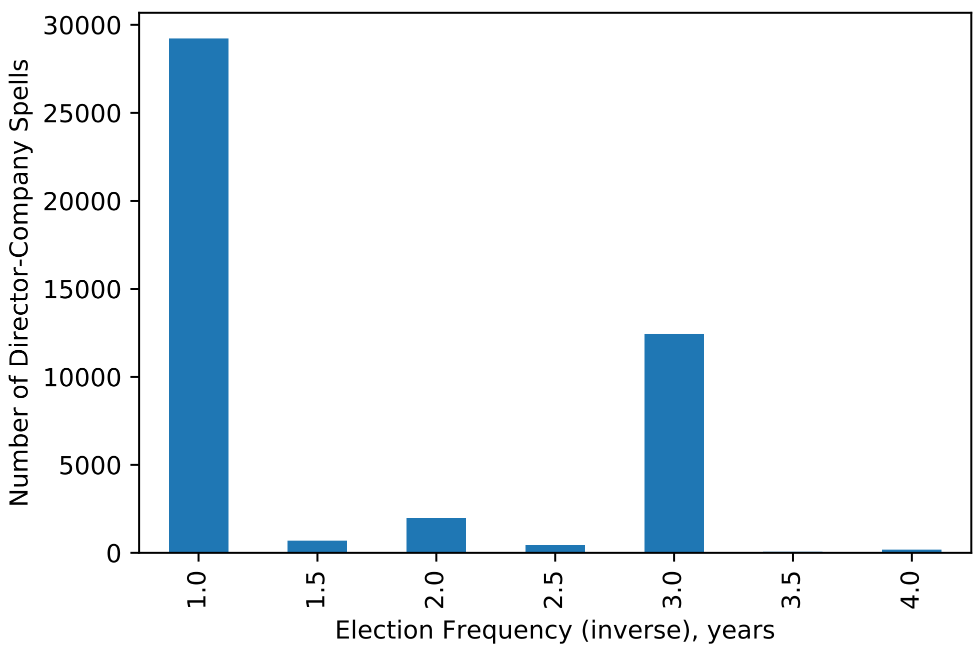

Figure 2.7: Election frequency (inverse) of director elections computed at the level of Company-Director spells. Years on horizontal axis represent the time between nominations of the same person for a director position at the same company. Election frequency of a spell is computed as the median of all elections' intervals belonging to the spell. The non-integer numbers correspond to spells having multiple different election intervals with no clear median.

The director elections data comes from the ISS Voting Analytics dataset. I merge it with the ISS Directors dataset to obtain directors' characteristics. Unfortunately, the match covers only about 40% of directors' election events. Therefore, I first rely on the Voting Analytics dataset to identify directors' employment spells. A detailed description of the procedure is available in the Appendix.

The Voting Analytics dataset contains 234305 director election events at 5616 companies for the years 2003 - 2016. This corresponds to 65788 identified director-company employment spells. Out of these, 19657 spells contain just a single election event. This is likely due to director being elected just once at a particular company.[16] 22437 spells were censored as the next expected election in a spell would not have been captured if it was to occur after 2016.

The frequency of director elections varies by company. A substantial part, 28% of director-company spells with two or more elections, of director elections happened at companies with staggered boards. This manifests in long waiting periods between nominations of the same person (typically, 3 years). Figure 2.7 presents the distribution of election frequencies for director-company spells. Election frequency of 2 years likely represents employment spells at companies transitioning between staggered and non-staggered boards of directors.

Directors' employment spells substantially differ in duration. Figure 2.8 shows the distribution of employment spells durations. As expected, shorter spells are more numerous in the data. Therefore, it is probable that director turnover is higher among board members with less experience at a company. Spells at companies with staggered boards cause 6, 9, and 12th bars to be noticeably higher than their neighbours. Figure 2.9 demonstrates that almost every election bears a risk to be the last one for a director at a company. As directors virtually never formally lose an election if nominated, the last election precedes a decision of leaving company's board of directors.

2.4.2 Election results and future nominations

Director elections do not seem to prevent nominees from getting on the boards yet their results do not go unnoticed. I find that low shareholder support predicts the event of director leaving the company. Director with higher shareholder support at the last election has a higher chance of participating in the next election. The magnitude of this relationship

[16] Some portion of these events might also be related to underrepresented companies in the ISS Voting Analytics dataset. As well as spells that ended in the first year (3 years) or started in the last year (3 years) of observable data.

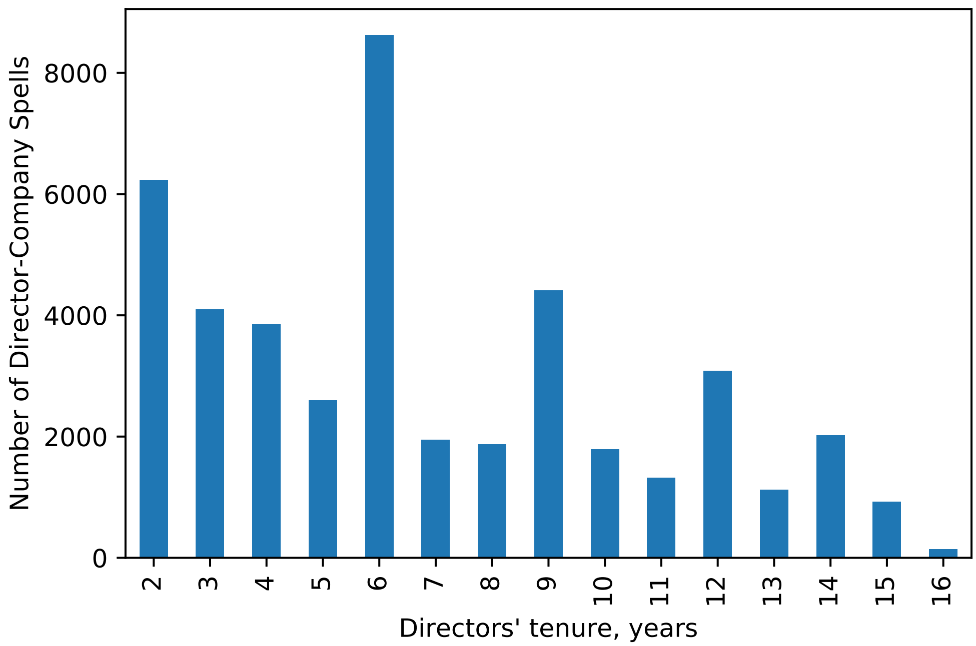

Figure 2.8: Tenure of directors' employment spells computed using ISS Voting Analytics data. Spell length is computed as a difference between the spells' maximum and minimum years of election plus the median time difference between elections in the spell. Due to the nature of procedure used, one-year-length spells are not identifiable as they only include one election. Three-year-spells are underestimated here due to the absence of 1-term directors from staggered boards in this statistic. Results in this bar chart are only representative of the sample used. Censoring occurs since only a limited timeframe is available for study. Actual directors' employment spells are likely to be longer.

is comparable to the effect associated with director being a member of compensation or audit committees.

To understand the effect of election results on director's tenure at a company, I consider a simple model of directors nomination. I suppose that directors and nomination committees make the decision regarding further nomination for a continuing director partially based on his/her last election results. That is, director may not want to re-elect, if he/she expects a weak support from shareholders. While this weak support is almost surely would be sufficient for a formal victory, a lower than his/her peers result might be a negative signal for the nominee's directorships at other firms. At the same time, nomination committee might have similar career concerns and, therefore, avoid nomination of unpopular directors.

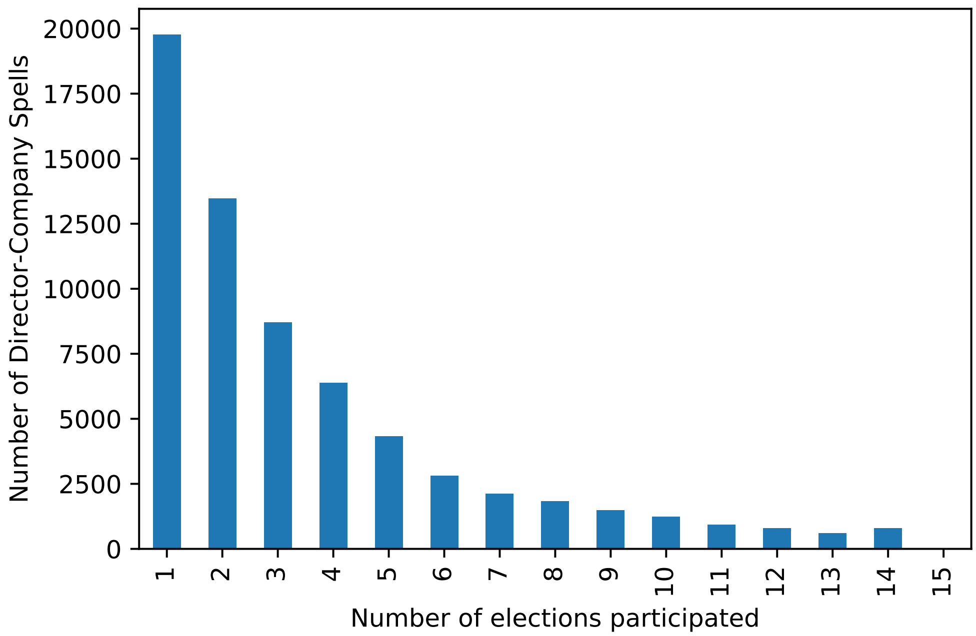

Figure 2.9: Number of elections a director participates in during his/her employment spell at a company. Bar chart represents results for the dataset in use. Due to censoring, actual numbers are likely to be higher.

Less popular nominees might also compromise board's defenses against activist shareholder's efforts to engage in a proxy fight.

For every director-company spell in the sample, I identify the last election a director participated in. I remove all censored spells from the sample because this simple model does not account for censoring. Using an election as a unit of measurement and within the framework of logit regression model, I regress an indicator of an election being the last one for a director on his/her performance in this election, ISS recommendation, and director's and company's characteristics. Table 2.6 presents the results.

The first, base model, specification only includes election results and a few election-specific characteristics. I consider the percentage of votes "For", support rate, among total outstanding voting shares as a measure of shareholder approval. Directors with higher support rates are more likely to be nominated in the next election cycle at a company. A one standard deviation increase in the support rate, 14 p.p., is associated with 21% increase in

Table 2.6: Directors abandoning future nominations for board elections at companies. Dependent variable is a dummy equal to 1 if election is the last one for a director candidate at a company. Censored employment spells are excluded from all samples. Support rate reflects the percent of votes "For" among total voting shares outstanding.

	Base model	Director controls	Company controls	Kitchen Sink
	(1)	*(2)*	*(3)*	*(4)*
Support Rate, %	-0.013^{***}	-0.010^{***}	-0.012^{***}	-0.012^{***}
	(0.001)	(0.004)	(0.001)	(0.004)
Staggered Board	1.632^{***}	2.845^{***}	1.595^{***}	2.822^{***}
	(0.087)	(0.274)	(0.101)	(0.308)
Majority Vote req.	0.602^{***}	0.998^{***}	0.438^{***}	1.045^{***}
	(0.080)	(0.172)	(0.093)	(0.203)
ISS "For" recommendation	-0.073	0.057	-0.016	0.011
	(0.070)	(0.230)	(0.085)	(0.272)
Support Rate × Staggered Board	-0.002	-0.010^{***}	-0.001	-0.010^{***}
	(0.001)	(0.003)	(0.001)	(0.004)
Support Rate × Maj. Vote. req.	0.004^{***}	-0.007^{***}	0.005^{***}	-0.007^{***}
	(0.001)	(0.002)	(0.001)	(0.003)
Support Rate × ISS "For" rec.	0.003^{***}	0.002	0.002	0.003
	(0.001)	(0.004)	(0.001)	(0.004)
log(Spell Length)	-1.311^{***}	-1.367^{***}	-1.346^{***}	-1.410^{***}
	(0.015)	(0.032)	(0.018)	(0.036)
Director's characteristics				
Director's Age		-0.102^{***}		-0.101^{***}
		(0.016)		(0.019)
(Director's Age)^2		0.001^{***}		0.001^{***}
		(0.000)		(0.000)
Director's Share, %		0.013^{***}		0.012^{**}
		(0.004)		(0.005)
Nominating Committee memb.		-0.057		-0.077
		(0.128)		(0.136)
Governance Committee memb.		0.035		0.050
		(0.128)		(0.136)

Table 2.6, continued

	(1)	*(2)*	*(3)*	*(4)*
Compensation Committee memb.		-0.143^{***}		-0.147^{***}
		(0.030)		(0.034)
Audit Committee memb.		-0.191^{***}		-0.186^{***}
		(0.030)		(0.034)
Employed as CEO		-0.164^{**}		-0.229^{***}
		(0.073)		(0.083)
Employed as VP		-0.630^{*}		-0.625^{*}
		(0.348)		(0.351)
Other employment controls	No	Yes	No	Yes
Company's characteristics				
log(Total assets)			0.039^{***}	-0.026^{***}
			(0.005)	(0.010)
Return on assets, %			-0.003^{***}	-0.005^{***}
			(0.000)	(0.002)
Book to market ratio			0.000	-0.060^{*}
			(0.000)	(0.032)
Leverage			-0.009	0.001
			(0.008)	(0.015)
Constant	1.585^{***}	3.749^{***}	1.309^{***}	4.020^{***}
	(0.068)	(0.567)	(0.087)	(0.652)
Observations	105196	30622	79589	24674

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

the odds of being nominated next time. This effect is comparable in magnitude to the effects of compensation or audit committees' membership, or director being employed as a CEO. In all four specifications I observe the same sign and similar magnitudes of this effect.

The staggered board dummy carries a substantial coefficient, but this is an artificial effect coming from the mechanics of staggered boards: a director, sitting on a board, does not participate in elections every year. This significantly reduces the number of nominations for these directors over their employment spell at a company in comparison to non-staggered boards' directors at other companies. Therefore, an election for a staggered board director is intrinsically more likely to be the last one. A similar outcome is observed for the majority voting requirement. As large number of companies transitioned from plurality to majority voting requirement over the sample's timespan, many director spells have ended under the new requirement while they were primarily lasting under the old requirement. Thus, the monotonous adoption of majority voting standard likely stands behind the significant positive coefficient here.

Surprisingly, a favourable ISS voting recommendation does not seem to have an effect on director's decision to participate in further company's elections.

Interaction terms of the support rate with the dummies described above produce mixed results. The support rate's effect is amplified at companies with staggered boards but only when controlling for director's characteristics. The interaction with the majority voting requirement has a similar dynamic: a higher support rate increases the chance of a future nomination when I control for the director's characteristics. The positive ISS recommendation interaction with the support rate does not have a significant effect anywhere except the base model specification where the effect is small.

Notably, the spell length has a substantial positive effect on the probability of participation in future elections. The longer a director stays on a board, keeping everything else constant, the greater chances are that he/she will be nominated at the next elections. This effect is present in all four specifications considered.

In the second specification I add director's characteristics to the base model. I find that the director's age has a "parabolic" relationship to the chances of future nomination. Younger and older directors have smaller chances of being nominated in comparison to middle-aged directors. The director's share in a company negatively affects his/her nomination chances, yet the effect is rather small as directors typically do not hold more than a fraction of a percent of company's shares. Membership in compensation and audit committees is associated with an ample increase in chances of future nomination. The same holds true for directors being employed as a CEO or a VP.

In the third specification I include company's characteristics to the base model. A higher return on assets seems to positively affect the director's chances of participation in future elections. The effect of the size of a company does not have a stable sign and depends on inclusion of director's characteristics. Other company's controls do not produce results that are consistently significant.

Finally, I use "kitchen sink" regression as my fourth specification. For most of the variables considered, I obtain the same signs and comparable magnitudes of coefficients. The variable of interest, director's support rate, remains significant and maintains its sign across all four specifications.

2.4.3 Election results and directors tenure

A low shareholder support is also associated with shorter director employment spells. In a survival analysis, I find that higher shareholder support is related to longer duration of director's employment at a company when I account for temporal evolution of the shareholder support level and the company's and director's characteristics. Notably, a survival analysis done in averages does not lead to a significant effect of shareholder support level on employment spell duration.

The survival analysis provides a capability to account for director-company spells that are censored due to the limited scope of the data sample.[17] This approach allows me to mitigate influence of possible biases that might have been introduced in the previous analysis by spell selection due to censoring. Since the survival analysis deals with employment spell's length as its main dependent variable, I adjust the hypothesis accordingly. The above analysis of decision to exit shows that a lower support rate is associated with a higher probability of not participating in future elections. A reasonable extension of this result could be a hypothesis that low support rates go in conjunction with reduction in the employment spell's length.

[17] More than 34% of spells in the sample are censored.

I use Cox's proportional hazard model to implement a survival regression. It features separation between the influence of static covariates and a population-level baseline hazard function in modeling of an individual's hazard. Moreover, this model allows me to estimate the effect of covariates without the need to estimate or assume a specific form of the baseline hazard function. As many covariates vary over the duration of a director-company spell, I also consider Cox's time varying proportional hazard model. Therefore, I study the relationship between support rates and employment lengths using the survival analysis in two settings: with static and time-varying covariates.

The analysis with static covariates involves the use of a statistic that maps values of a set of time-varying variables into their static counterparts. I chose to use the mean as such statistic for all variables. In an unreported analysis, I find that the mean support rate of a director at a company is either not significantly related to his/her employment spell's length or the magnitude of such relationship is not economically meaningful.[18] In this analysis, I consider regression specifications that include director and company controls, as well as, all and none of those. In addition, I construct other statistics that summarize director's support: the mean excess support in comparison to his/her peers, the last support rate, and the last excess support in comparison to historic performance of the director herself/himself.[19] In all but one case, I do not find a significant and meaningful effect on director-company's spell length.[20]

The lack of results in a static covariates approach in this case is likely related to the inability of simple statistics to convey meaningful information about the director's election performance. For instance, the average support rate might not be easily comparable across

[18] I find the mean support rate as an average of all positive support rates of a director at all elections he/she participated in at the company.

[19] The mean excess support in comparison to director's peers is computed as a difference between the director's support and the mean directors' support in a given election year averaged over all election years the director participated in. The last support rate corresponds to the support rate of the director in the last election she/he participated in at the company. The last excess support rate is computed as a difference between the director's last support rate and his/her average support rate.

[20] A significant and economically meaningful effect has been found for the last excess support rate variable in specification with controls for director's and company's characteristics.

the companies and employment spells that are too distant in time. At the same time, averaging of company's and director's time variant characteristics severely reduces their informativeness in the regression. That is, for example, a director's age becomes less relevant for longer spells, a company's return on assets means much less for career prospects of a specific director when averaged over many years, and within spell covariation between explanatory variables vanishes after averaging.

To address the repetitive nature of elections in company-director spells, I utilize Cox's time varying proportional hazard model. Formula 2.1 describes how time changing covariates are embedded into the model.

$$h(t|x) = b_0(t) \exp\left(\sum_{i=1}^{n} \beta_i (x_i(t) - \overline{x_i})\right), \tag{2.1}$$

where $h(t|x)$ is the conditional hazard function, $b_0(t)$ is the baseline hazard, $x_i(t)$ are the time-varying covariates, and β_i are the survival regression coefficients. For every director-company spell, I assume that covariates are updated at the time of the director's election and then they stay the same until the next election of this director happens.

The analysis with time-varying covariates provides evidence that a higher support rate is linked to longer duration of director-company spells. Table 2.7 demonstrates results of four different specifications: base model, models with director's and company's controls, and a "kitchen-sink" model. The support rate has a significant coefficient in all four specifications, and it has a substantial magnitude in all but the base model specification. Among 3 other specifications, an increase of one standard deviation in the support rate, is associated with more than a 6% decline in the hazard rate. When coupled with a staggered board, the effect grows to a 14% decline. A positive ISS voting recommendation and a majority voting requirement contribute an additional 2% each to this support rate effect.

A staggered board effect measures at around 38% reduction in the hazard rate. This is likely due to a large number of directorships at non-staggered boards that do not last longer

Table 2.7: Survival analysis of director-company spells' duration using Cox's time varying proportional hazard model. The table presents estimates of β_i coefficients for the following model of conditional hazard function: $h(t|x) = b_0(t)\exp\left(\sum_{i=1}^{n}\beta_i(x_i(t) - \overline{x_i})\right)$, where $b_0(t)$ is the baseline hazard and $x_i(t)$ are the time-varying covariates.

	Base model	Director controls	Company controls	Kitchen Sink
	(1)	*(2)*	*(3)*	*(4)*
Support Rate, 100%	-0.022^{***}	-0.477^{***}	-0.301^{***}	-0.551^{***}
	(0.004)	(0.096)	(0.025)	(0.110)
Staggered Board	-0.464^{***}	-0.467^{***}	-0.534^{***}	-0.493^{***}
	(0.013)	(0.042)	(0.016)	(0.047)
Majority Vote req.	-0.082^{***}	-0.106^{***}	-0.066^{***}	-0.072^{**}
	(0.012)	(0.028)	(0.015)	(0.032)
ISS "For" recommendation	-0.118^{***}	-0.039	0.001	-0.018
	(0.012)	(0.063)	(0.017)	(0.071)
Support Rate × Staggered Board	-0.551^{***}	-0.573^{***}	-0.534^{***}	-0.599^{***}
	(0.017)	(0.053)	(0.020)	(0.059)
Support Rate × Maj. Vote. req.	-0.223^{***}	-0.160^{***}	-0.164^{***}	-0.133^{***}
	(0.016)	(0.036)	(0.019)	(0.041)
Support Rate × ISS "For" rec.	-0.020^{***}	-0.197^{***}	-0.130^{***}	-0.197^{**}
	(0.004)	(0.072)	(0.020)	(0.082)
Director's characteristics				
Director's Age / 100		0.388^{**}		0.458^{**}
		(0.180)		(0.208)
(Director's Age)^2 / 10000		0.514^{***}		0.567^{***}
		(0.152)		(0.175)
Director's Share, 100%		-0.008		-0.006
		(0.010)		(0.011)
Nominating Committee memb.		-0.014		-0.015
		(0.029)		(0.033)
Governance Committee memb.		-0.020		-0.020
		(0.029)		(0.033)
Compensation Committee memb.		-0.040		-0.030
		(0.025)		(0.029)

Table 2.7, continued

	(1)	*(2)*	*(3)*	*(4)*
Audit Committee memb.		-0.102^{***}		-0.098^{***}
		(0.024)		(0.028)
Director's employment controls	No	Yes	No	Yes
Company's characteristics				
log(Total assets)			-0.069^{***}	-0.086^{***}
			(0.003)	(0.008)
Return on assets, %			-0.003^{***}	-0.005^{***}
			(0.000)	(0.001)
Book to market ratio			0.063^{***}	0.025
			(0.005)	(0.028)
Leverage			0.041^{***}	0.009
			(0.004)	(0.014)
Observations	222830	33859	149875	26061

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

than 1 or 2 terms. Therefore, being on a staggered board even for a single term delivers a sizeable impact on the length of director's tenure. The majority voting requirement is also associated with longer directorships with the effect being in the neighbourhood of 8%.

A positive ISS recommendation does not seem to have a substantial effect on its own. Only in the base model it delivers a 11% decrease in the hazard function, while in other specifications the effect loses its significance.

The director's age, in this analysis, has a negative influence on employment spell duration. A 10-years change in age translates into a 4.5% increase in hazard function. Unlike in the nomination analysis above, a compensation committee membership does not have a significant effect on the spell's length. An audit committee membership has a positive effect on the spell's duration as before.

Directors at larger companies enjoy longer employment spells. While a higher return on assets increases duration of directorships, the effect is rather small: less than a half-percent decrease in hazard function per a percent increase in the return on assets. Higher book-to-market and leverage ratios reduce length of employment spells, but this effect only persists in the company's controls specification and disappears in the "kitchen-sink" regression.

2.5 Portfolio composition and voting behavior

In a world where firms are price takers, Fisher separation theorem (Fisher, 1930) establishes that all shareholders, no matter what portfolios they hold, should unanimously agree on the firm's production plan that maximizes its profit. In the real world, price taking assumption is unlikely to be satisfied, and we may observe a shareholder behavior that is not compatible with the firm's profit maximization objective.

In this section, I study the effect of mutual funds mergers on voting behavior of acquiring funds. In particular, in an event study I demonstrate that a merger with another fund causes a noticeable change in how the acquiring fund votes the shares it holds. Since a merger is likely associated with a change in the acquiring fund's portfolio, this study suggests that there might be an effect of portfolio structure on the fund's voting behavior. At the same time, mergers may also lead to other adjustments for the acquiring fund that might cause the change in its voting behavior.

Portfolio endogeneity presents a substantial hurdle in the analysis of shareholder's voting behavior. Since shareholders may vote both with their shares and with their feet (Admati & Pfleiderer, 2009), the direct comparison between voting behavior and portfolio structure might produce spurious results. For example, investors, like mutual funds, may follow certain sets of principles to select assets into their portfolios and to vote their shares. Therefore, groups of mutual funds having similar principles could create a correlation between their portfolio structure and their voting record.

My analysis builds on the assumption that the reasons for a merger of mutual funds are not directly related to their voting behavior. Literature finds that a poor target fund performance is a significant factor for within-family mergers. Jayaraman et al. (2002) find that eliminating funds with high cost structures and disguising poor fund performance are the likely reasons for within-family mergers, while building a larger set of investment objectives is a probable goal for across-family mergers. Fund families are likely to sell unique portfolios to other mutual fund families in order to stay focused (Zhao, 2005). McLemore (2019) finds that fund's past performance is not significantly related to the likelihood of it being an acquiring fund. Khorana et al. (2007) find that when a target fund's board has many independent directors the chances of a merger for an underperforming fund are higher. I have not been able to find studies that cover voting behavior of merging mutual funds.

2.5.1 Mutual funds mergers and voting data

The mutual funds' mergers data comes from the CRSP Mutual Fund database. I find 1346 fund mergers that happened from 2009 to 2016.[21] For each acquired fund the database provides the date of a merger and the acquirer information. I use it to construct the set of funds that survived the merger (acquirers) and experienced a shock to their portfolio.

The mutual funds' votes come from the ISS Voting Analytics database. For every acquiring fund I collect all its votes within a two-year timespan (one year prior and one year post merger). The voting behavior analysis involves learning the differences between votes of the fund in question in comparison to votes of the other funds. Thus, I also collect the votes of other mutual funds at the meetings where the acquiring fund was actively present.

2.5.2 Synthetic control method

I analyze an acquiring fund's voting behavior by comparing it to the behavior of other funds that vote at the same companies. This poses a challenge as mutual funds are different from each other, hold non-identical portfolios, and typically do not exhibit identical voting patterns. I find that synthetic control approach fits the problem well, and I use it to construct

[21] My choice for the time interval is explained by the mapping I created between CRSP Mutual Fund and ISS Voting Analytics databases. While the mapping works for a greater timespan, from 2009 to 2016 the match is substantially better than for other years.

an artificial, "synthetic", mutual fund that tracks the voting behavior of the acquiring fund before a merger (treatment). This allows me to reconstruct a counterfactual case where the would-be acquiring fund is not treated by a merger. Then, I use this synthetic control to find a difference in voting behavior of the acquiring fund after the treatment in comparison to the counterfactual case.

Abadie et al. (2010) introduce synthetic control method in their study of California's tobacco control program. I also adopt the factor model they propose as I find it suitable for a study of funds' voting behavior. In particular, let Y_{it}^N be the voting decision of a non-treated fund-i at an election enumerated by a timestamp t. Consider the following factor model

$$Y_{it}^N = \delta_t + \boldsymbol{\theta}_t \boldsymbol{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \epsilon_{it}, \tag{2.2}$$

where δ_t is an unknown common time-varying factor affecting funds' votes, $\boldsymbol{Z}_i$ is a vector of observed time-independent covariates, $\boldsymbol{\theta}_t$ is a vector of unknown parameters, $\boldsymbol{\lambda}_t$ is a vector of unobserved common time-dependent factors, $\boldsymbol{\mu}_i$ is a corresponding vector of unknown factor loadings, and ϵ_{it} represents the error term with zero mean.

In this paper, I interpret time as an indexing axis for the election events a fund participates in.[22] Thereby, I can come up with the following rationalization for the variables involved. The vector $\boldsymbol{Z}_i$ represents the observed characteristics of a mutual fund that do not change with time and are not affected by a merger.[23] For example, these can be fee structure, published investment strategy, fund's management and the board of directors. To account for a differential impact of fund's covariates on its voting decision at a particular election, vector $\boldsymbol{\theta}_t$ contains unknown weights that apply to the covariates in $\boldsymbol{Z}_i$. Since $\boldsymbol{\theta}_t$ is time-dependent, these weights can be election-specific which allows for a great deal of flexibility in accounting for fund-specific covariates' effect on the fund's voting behavior. In

[22] Without loss of generality, I assume that every election issue can be assigned a "time" that uniquely identifies it. Then election and company covariates can be embedded into the "time-dependent" variables.

[23] Since I track acquiring fund's behavior only within a fixed time window of 2 years, slow changing characteristics of a mutual fund can be treated as time-invariant within this model's framework.

a similar fashion, vector $\boldsymbol{\lambda}_t$ contains unobserved election-specific characteristics that affect fund's voting decision. A corresponding vector of unknown weights, $\boldsymbol{\mu}_i$, reflects how mutual funds are taking into account those election-specific covariates. Finally, variable δ_t takes care of election-specific effects that uniformly affect funds' voting decisions.

The synthetic control method uses a pool of J donors, mutual funds not involved in a merger ($i = 2, ..., J+1$), to construct an estimate of the counterfactual outcome, $\hat{Y}_{1t}^N = \sum_{j=2}^{J+1} w_j Y_{jt}^N$, for the treated unit, the acquiring mutual fund ($i = 1$). The construction involves a set of weights, $w_2, ..., w_{J+1}$, that are tuned in order to match the pre-intervention voting path (2.3) of the acquiring fund and its observed covariates (2.4) as close as possible.

$$\sum_{j=2}^{J+1} w_j Y_{j1} = Y_{11}, \quad \sum_{j=2}^{J+1} w_j Y_{j2} = Y_{12}, \quad ... \quad \sum_{j=2}^{J+1} w_j Y_{jT_0} = Y_{1T_0}, \tag{2.3}$$

$$\sum_{j=2}^{J+1} w_j \boldsymbol{Z}_j = \boldsymbol{Z}_1, \tag{2.4}$$

where T_0 is the last election before the merger.[24] It may not always be feasible to find a set of weights such that the sets of equations 2.3 and 2.4 hold exactly. In such cases, an approximate solution is sought. The synthetic control method does not specify how to find a tradeoff between better approximation of one system of equations over the other and vice-versa. Since in this section I concentrate on a qualitative study of voting behavior, I focus solely on the pre-merger voting path match and I ignore matching on funds' characteristics when searching for the synthetic control's weights.

The benefit of using the synthetic control method over a diff-in-diff regression or a similar model is the absence of the parallel trend assumption. Abadie et al. (2010) show that weights $w_2, ..., w_{J+1}$ can only fit (approximately) the systems of equations 2.3 and 2.4 with a high number of pre-intervention periods if these weights approximate $\boldsymbol{\mu}_1$ through a weighed sum of $\boldsymbol{\mu}_j$, $j \in \{2, ..., J+1\}$. That is, if there is a non-linear trend in voting behavior of a mutual fund

[24] Here, for the sake of notational simplicity, I assume that elections are enumerated by integers that represent a time sequence.

of interest, this trend will be picked up by non-linearities in behavior of mutual funds in the donor pool. Therefore, the synthetic control method accounts for the influence of unobserved election-specific characteristics in the estimate of acquiring fund's voting behavior in the counterfactual case.

One drawback of the synthetic control method is the requirement that every fund in the donor pool has to have a voting history that completely covers all votes of the acquiring fund. Since mutual fund portfolios almost never overlap exactly, an additional step is needed before I can apply the method.

2.5.3 Robust synthetic control method

To overcome the problem of missing data in the donor pool of mutual funds, I use the robust synthetic control method developed by Amjad et al. (2018). The idea behind this method is to perform a spectral decomposition of funds' voting histories and then inverse this procedure to impute missing voting data. This way I can reconstruct the would-be votes of a mutual fund from the donor pool at companies that are not in its portfolio.

The spectral decomposition relies on a well-balanced sample of donor funds. For every acquiring fund in the sample, I construct a donor sample and conduct a separate robust synthetic control analysis. At the first step, I determine all elections within a two-year span the acquiring fund participated in. Then, I collect votes of all other funds that have participated in any of those elections. At the second step, I prune the set of the other funds participated by leaving only those that have voted in at least 50% of elections before and after the merger date. Next, I remove acquiring funds with donor pools of fewer than 4 funds from consideration. Figure 2.10 presents the number of funds in a donor pool and the number of elections considered for the merger events at acquiring funds. For the majority of mergers, the voting path consists of more than a thousand votes cast by the acquiring fund and the corresponding donor pool contains more than ten donors.

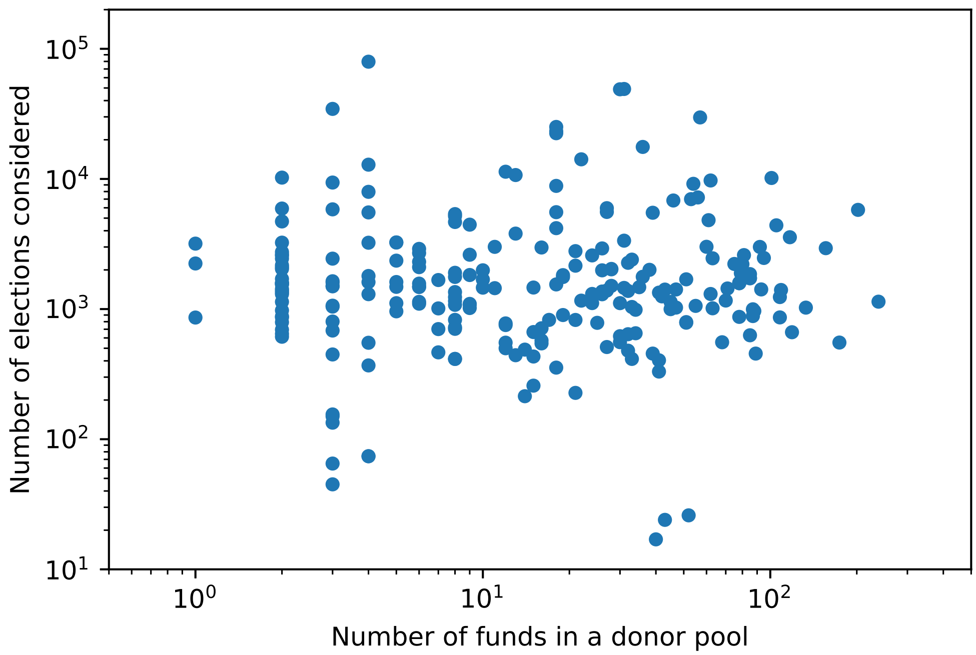

Figure 2.10: Sizes of donor pools and considered vote-paths for the acquiring funds' merger events. Vote-paths are tracked based on the elections an acquiring fund participated in over a two-year window around the merger date. Funds in the donor pool are required to have votes in at least 50% of both pre- and post-merger elections.

2.5.4 Merger's effect on voting behavior

Acquiring mutual funds experience a significant change in their voting behavior right after a merger. I track changes in voting behavior of a mutual fund by comparing it to an implied behavior of such fund in a counterfactual case where the merger never happens. In particular, using the weights computed by robust synthetic control method, I compute the implied voting path of the acquiring fund.

$$\hat{Y}_{1t}^{N} = \sum_{j=2}^{J+1} w_j \hat{Y}_{jt}^{N}, \tag{2.5}$$

where $\hat{Y}_{jt}^{N}$ is an estimate of numerically encoded vote of fund-j at election-t. $\hat{Y}_{jt}^{N}$ is a product of the robust synthetic control method and is based on the observed part of the fund-j's voting-path, Y_{jt}^{N}. I set $Y_{jt}^{N} = 1$, if fund-j voted "For", $Y_{jt}^{N} = -1$, if the fund voted "Against", "Abstain", or "Withhold", and I use $Y_{jt}^{N} = 0$ to reflect a recorded "Do Not Vote". If I can not find a recorded vote of the fund-j at election-t, I leave Y_{jt}^{N} as missing value for the spectral decomposition to impute.

I compute the difference in voting behavior between an acquiring mutual fund's voting-path and the counterfactual case as an absolute difference between the encoded vote of the acquiring fund and its implied voting-path. I adjust the computed difference by a factor of $\frac{1}{2}$, so it can be interpreted as a share of cases in disagreement on a scale from 0 to 1. The resulting vector of differences turns out to be very noisy for any individual acquiring mutual fund. I group the computed differences into weekly intervals by the elections' dates. Then, I use a weighted average to compute a single measure of voting disagreement within a weekly period along with a 95%-confidence interval. The weights are inversely proportional to the number of election events an acquiring fund participated in. Weighting scheme has a purpose of preventing the few funds with high number of votes cast from skewing the results of averaging.

Figure 2.11 presents a clear jump in voting behavior disagreement between acquiring mutual funds and their corresponding synthetic controls. In the weeks before a merger, the disagreement level measures at around 8%, while by the end of the first few weeks after the merger it spikes to 14% and then consolidates at the level of 11%. It is hard to judge what happens after a few months from the merger, as long-term predictions with synthetic control method are less reliable when pre-intervention histories are short.[25]

2.5.5 A placebo study

The synthetic control method may only use the pre-merger voting path to construct the control's weights. This raises a credible concern that if overfitting happens then we can see a spike in voting behavior disagreement just because the synthetic control performs badly

[25] While some merger events have very extensive pre-merger voting-paths, others have limited numbers of acquiring fund's votes recorded. Averaging of a heterogeneous set of post-intervention estimates likely leads to worse estimates at long time-horizons as the weighting scheme prefers estimates with smaller number of votes and, consequently, shorter histories.

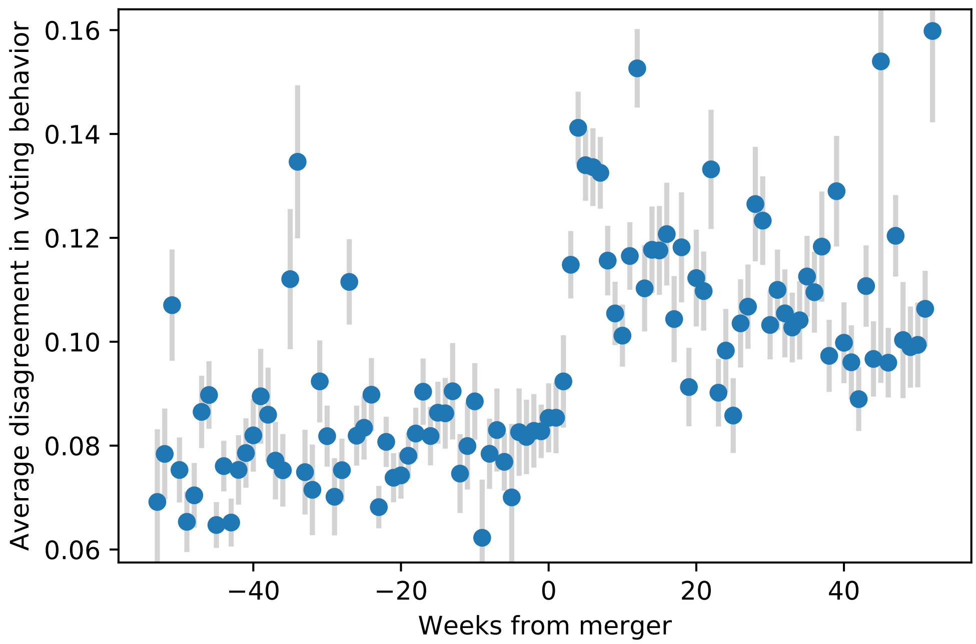

Figure 2.11: The average disagreement in voting behavior between acquiring funds and their corresponding synthetic controls. The disagreement is computed on a scale $[0, 1]$ as a half of an absolute difference between encoded fund's and its control votes. Averaging involves a weighting scheme to prevent acquiring funds with extensive vote-paths from dominating the results. Weights are inversely proportional to the number of votes an acquiring funds has. 95% confidence intervals are depicted in gray and are computed separately for each week's value.

on the unseen data. Another problem could be caused by timing of mutual funds' mergers. I observed a substantial heterogeneity in placement of funds' merger dates with respect to the dates of shareholder meetings. This heterogeneity could lead to an artificial jump in disagreement simply due to a merger being scheduled, for instance, just before a month with the highest number of shareholder meetings.

To address the concerns above, I implement a placebo study where I replace acquiring mutual funds with arbitrary funds that did not experience mergers. The nature of the study allows me to include more data points than the original study could by considering more than one arbitrary fund per one merger date. This substantially reduces variance in the resulting graph. Figure 2.12 presents results of the placebo study.

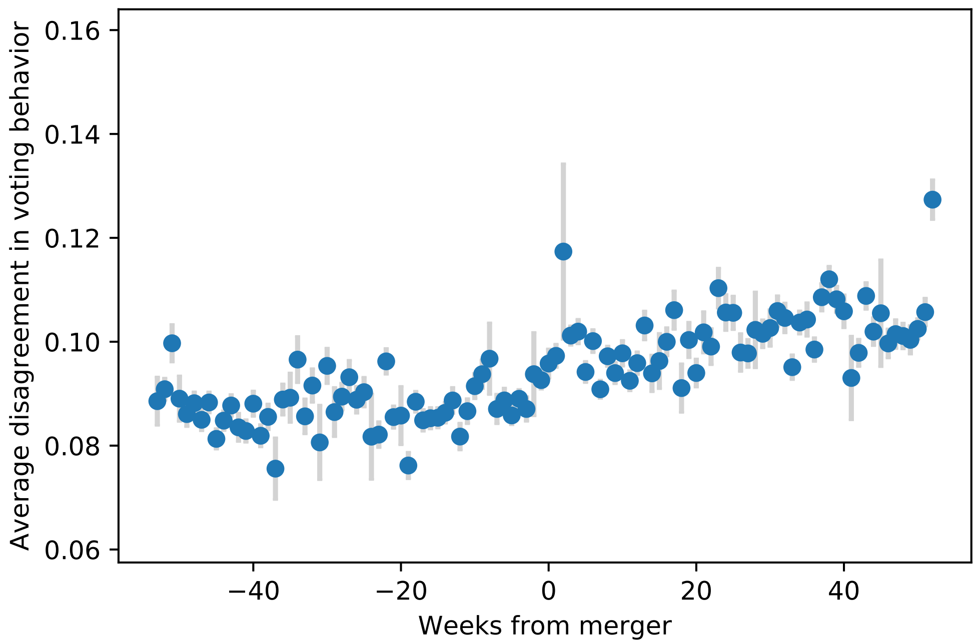

Figure 2.12: A placebo study of the average disagreement in voting behavior of mutual funds after mergers. The study involves the actual merger dates, but the acquiring funds are replaced by arbitrary non-merging funds. The disagreement is computed on a scale $[0, 1]$ as half of an absolute difference between encoded fund's and its control votes. Averaging involves a weighting scheme to prevent the funds with extensive vote-paths from dominating the results. Weights are inversely proportional to the number of votes a funds has. 95% confidence intervals are depicted in gray and are computed separately for each week's value. To reduce variance of the computed values, multiple arbitrary funds were considered per one merger date.

The study shows no jump at a merger date. This result strengthens the validity of the jump in the actual study. A notable feature of placebo test is an upward trend for the post-merger disagreement values. This trend is likely a result of the drift in funds' portfolio structures that happens over time. Another reason could be the method's limited ability to predict the counterfactual case's outcome at longer time intervals.

2.6 Conclusion

In my analysis of shareholder meetings, I find that director elections and auditor ratifications appear at more than 94% and 80% of shareholder meetings respectively, and together they comprise more than 80% of all recorded election events in the ISS Voting Analytics dataset. Using principal component analysis, I establish that the presence of director elections on shareholder meeting's agenda is one of the least significant factors that distinguish between different compositions of shareholder meetings. This makes director elections a good candidate for further study as those almost uniformly appear across shareholder meetings. Then, I find that the two most frequent election issues, director elections and auditor ratifications, have skewed distributions of election tallies. Less than 4.7% of director elections and less than 1% of auditor ratifications received support from less than half of shares outstanding.

For uncontested director elections, I test two hypothesis of delayed effects of election results on a director's career prospects. First, I find that low shareholder support predicts an event of director leaving the company. This result is coherent with Aggarwal et al. (2019) who find a similar result for a greater percentage of shareholder dissent votes. Second, I discover that low shareholder support is associated with shorter director-company spells. This result is comparable in magnitude to the effect of an audit committee membership.

Finally, in an event study, I find that a merger with another mutual fund causes the acquiring fund to change its voting behavior. Since mergers are likely to modify portfolio structure, this study suggests that portfolio composition affects voting behavior. At the same time, mergers may introduce many other modifications to the acquiring fund that might be responsible for the change in its voting behavior.

Appendix B

Director spells

An important step in studying the effect of election outcome on director turnover is identifying director employment spells. Unfortunately, the ISS Directors dataset covers only about 40% of firm-director pairs in the ISS Voting Analytics aggregate election results dataset. At the same time, the aggregate election results dataset provides a detailed description of agenda items that in the case of director elections includes the full name of director nominee. Therefore, I implement director's name matching procedure within each company to identify his/her spell.

At the first step, I use regex to extract directors' names from election agenda descriptions.[26] I also remove titles that follow the names, e.g. Ph.D., M.D., etc.

At the second step, I focus on each company separately to identify which director elections across the years correspond to the same person. I assume that no director was a nominee more than once per year,[27] and that there might be gaps between the elections of the same person.[28] While typically a person appears under an exactly identical name, some times there might be deviations in spelling and auxiliary names (nicknames) attached to his/her name. To overcome this issue, I consider the matching problem as multidimensional clusterization problem that allows for a some amount of noise to be present.

[26] Agenda descriptions for director elections are very well standardized. To extract directors' names in 99.98% of cases a regex expression with only 4 starting statements was required. The expression was `"(?:Elect +Directors{0,1}|Elect|Reelect|Director) +([\W \w ]+?)(?: as[\W \w ]+)*$"`.

[27] An exception from one-election-per-year assumption are elections with "Pending" vote result status.

[28] For example, some companies have staggered boards of directors. Thus, a director might be elected for a few years and will not appear in the next year's nominee list.

I use Levenshtein distance to compute the difference (distance) between the names in the pool. A distinctive property of this measure, unlike other string distance measures, is that it is a metric distance. That is, it satisfies the triangle inequality which in turn, allows me to "place" directors' names in a multidimensional space to perform a cluster analysis.

For cluster analysis I use Ordering Points To Identify the Clustering Structure (OPTICS) algorithm. It is closely related to a better known Density-based spatial clustering of applications with noise (DBSCAN) algorithm. Unlike DBSCAN, OPTICS is better suited for finding clusters in data of varying density. Directors' names in a pool can be represented by a set of points in a metric space with help of multidimensional scaling. The OPTICS algorithm implementation in Python allows me to directly use a pre-computed distance matrix which removes the need of using multidimensional scaling. To prevent spurious clusters from appearing in companies with just few director elections, I limit the maximum Levenshtein distance to 7 between two samples for them to be considered being in the same neighborhood.

Cluster analysis produces a two-part result. The first part is list of sets of directors' names where each set corresponds to one person based by his/her name similarity. The second part is a list of names that were difficult to match with any particular director's names set.

I use the following post-processing procedure in order to improve on the result of cluster analysis. First, I analyze the minimal Levenshtein distance between the name sets in the first part. If two cluster-identified directors have very small (less than 3) distance between the name sets and no overlap in the election years I join them together in one set and treat as a single person. This corrects for clustering algorithm behavior that multiple exact repetitions of a director's name lead to a wrongful rejection of a sightly different spelling of the same person's name. Second, I loop over the names in the unmatched set and see if these can be attributed to an already identified person. If no association is possible, I designate a new director persona for such name.

Chapter 3

Funds of a Feather: Influencing Corporate Elections by Voting Together

3.1 Introduction

The dramatic rise of institutional equity ownership over the last few decades warrants greater scrutiny over the effects it may have on corporate governance of the U.S. public firms.[1] Close attention in the literature is placed on the role of the "Big Three" (BlackRock, Vanguard, and State Street) as they command substantial shares in the largest U.S. companies and frequently populate the lists of top beneficiaries of respective companies.[2] Studies have shown that individual mutual funds within a mutual fund family tend to exhibit similar voting behavior at corporate elections.[3] By voting in a lockstep, a group of shareholders might exercise greater influence on company's governance. While literature views mutual fund families as blockholders, much less attention is devoted to study correlated voting

[1] Backus et al. (2019b); Baig et al. (2018)
[2] Bebchuk & Hirst (2019); Coates (2018)
[3] Fichtner et al. (2017)

behavior among non-related investors, e.g., individual mutual funds belonging to different families.

In this article, I explore the factors that are associated with higher chances that a pair of shareholders makes the same voting decision at corporate elections. I do this by observing votes of individual mutual funds at director elections. Then, I study how one such factor, similarity of shareholders' portfolios, affects shareholder participation.

I find a positive relationship between portfolio similarity of a pair of mutual funds and probability of their voting decisions being the same. I show that greater portfolio similarity among mutual funds leads to lower participation of other shareholders in director elections.

Theoretical literature provides a classical result, Fisher separation theorem (Fisher, 1930), that shareholders with heterogeneous portfolios should unanimously agree on actions that maximize a firm's profit under the necessary assumption that firms are price-takers (Milne, 1974; Hart, 1979; DeAngelo, 1981). Therefore, we should not observe heterogeneity in shareholders' voting decisions that is based on differences in their portfolios. The observed correlation between portfolio structure and voting decisions suggests that the price taking assumption is most likely violated (and shareholders no longer unanimously want to maximize the firm's profits).[4] In the absence of perfect competition, profit maximization is less likely to be a firm's objective (Hansen & Lott, 1996; Gordon, 2003). Contemporaneous literature suggests that maximization of a weighted sum of profits of shareholders' portfolios might be a reasonable alternative objective (Salop & O'Brien (2000); Azar (2017); Brito et al. (2018), see also Schmalz (2018) for a detailed review). Thus, a shareholder who wants to maximize the value of her portfolio may want to account for the effects of between-firm externalities when setting the firm's objective through voting at corporate elections.

Mutual funds' investment advisers have a fiduciary duty to vote proxies "in a manner consistent with the best interest of the fund and its shareholders" (SEC, 2003a). This ensures

[4]The other possible explanations might include limited ability of shareholders to collect and process information, thus differences in portfolio structures may correlate with differences in opinions of what is best for profit maximization.

high turnout by mutual funds at corporate elections and, in conjunction with imperfect competition, provides a testable hypothesis about their voting patterns. Under imperfect competition governance decisions at one firm may affect the financial outcomes of other firms. Therefore, a mutual fund, which strives to maximize its portfolio profit,[5] must internalize the effect of a voting outcome at one firm on the value of other firms in its portfolio. I reject the hypothesis that mutual funds' voting decisions are not related to their portfolios by observing a positive correlation between portfolio similarity and voting decisions.

This work contributes to the literature in three main ways. First, I study the voting behavior of individual mutual funds. Iliev & Lowry (2015) find that funds, that have higher net benefits of voting, more often vote independently of Institutional Shareholder Services (ISS) recommendation.[6] Fichtner et al. (2017) demonstrate that the "Big Three" families of mutual funds utilize coordinated voting strategies. Schwartz-Ziv & Wermers (2019) observe that institutional investors, when making voting decisions, account for firm's weight in their portfolio and their fraction-of-company investments. In contrast to the existing literature, I study how differences in characteristics of mutual funds affect the probability of them making the same voting decisions. I find that a mutual fund's family has a significant impact on the fund's voting behavior. This result goes in line with Fichtner et al. (2017), as funds from the same family tend to agree on voting decisions. I then discover that funds with more similar portfolios tend to cast identical votes more often. This finding provides evidence that individual mutual fund's portfolio structure is taken into account by the decision-making body. Unlike Schwartz-Ziv & Wermers (2019), I find that portfolio structure at the individual fund level explains the fund's voting behavior better than portfolio

[5]In a more realistic scenario a mutual fund would want to maximize profit of its portfolio subject to a variance constraint. Since between-firm externalities shape the joint distribution of fund's holdings' payoffs, the fund should internalize the effect of voting outcome at one firm on this joint distribution of payoffs. This would require mutual fund to account for its entire portfolio composition when voting at a single firm.

[6]Institutional Shareholder Services (ISS) is the largest proxy advisory firm. It routinely issues voting recommendations regarding how investors should vote on corporate questions.

structure of its mutual fund family.[7] I also find that favorable ISS recommendation reduces chances of disagreement between the funds. This result adds to the literature of mutual funds' reliance on Institutional Shareholder Services (ISS) and management recommendations on voting decisions (Choi et al., 2009, 2010; Iliev & Lowry, 2015; Malenko & Shen, 2016; Malenko & Malenko, 2019).

Second, I investigate how highly diversified horizontal shareholders[8] affect a company's governance process. My analysis shows that portfolio structure affects both: individual mutual funds' voting decisions and participation at directors elections as a whole. This contributes to the literature on the effects of horizontal shareholding and cross-ownership (Backus et al., 2019a; Elhauge, 2019a,b,c; Morton & Hovenkamp, 2018; Brito et al., 2019; He et al., 2019). I establish that higher portfolio similarity of mutual funds causes lower turnout at director elections by other shareholders. This adds to the literature on rational apathy of investors (Jill E. Fisch, 2017; Nili & Kastiel, 2016), network effects on voting (Enriques & Romano, 2018), and shareholder free-riding (Lafarre, 2017; Cvijanovic et al., 2019). Third, I extend the cosine portfolio similarity measure (Bohlin & Rosvall, 2014; Sias et al., 2013; Getmansky et al., 2018; Backus et al., 2019b) to evaluate portfolio similarity within sets of more than two shareholders.

This paper also provides a bridge between the literature examining the growth of large index fund families (Bebchuk & Hirst, 2019; Coates, 2018) and the literature on horizontal shareholding (Elhauge, 2019a,b,c; Morton & Hovenkamp, 2018; Brito et al., 2019; He et al., 2019). The former focuses on the power of small groups of exceptionally large mutual funds families in influencing corporate decision making, while the latter considers investors, not necessarily large ones, that hold multiple competitors in the same product market simultaneously. I observe that a mutual fund family is not the only source of power centralization.

[7]In an unreported regression, I study how similarity of mutual fund families' portfolios affect the probability that funds from different families make identical voting decisions. After conditioning on the same mutual funds' and their families' characteristics I do not find a statistically significant relationship.

[8]Elhauge (2019a) defines horizontal shareholding as an overlap of leading shareholders of horizontal competitors.

Since an increase in portfolio similarity tends to correlate with probability of different individual mutual funds making the same voting decisions, I infer that the boundaries between different mutual fund families might be blurred by the disperse nature of their funds' portfolios. Therefore, not only the number and size of the top mutual fund families matter for concentration of decision making power, but also their, and other shareholders, degree of portfolio similarity, which is driven up by diversification and horizontal shareholding.[9]

Free-riding is a possible explanation for lower turnout of shareholders at elections with high portfolio similarity among mutual funds. Retail shareholders may expect mutual funds votes to be aligned with their position (Cvijanovic et al., 2019); as well as shareholders may decide not to vote if they are less informed and want more informed voters to participate instead (Feddersen & Pesendorfer, 1996).

Alternative possible channels might include shareholders' rational apathy. In the previous chapter, I establish a firm's objective function and find that a firm pays greater attention to shareholders with correlated voting decisions than to the ones with no or negative correlation. Then, a reasonable explanation for the causal effect of portfolio similarity on shareholder turnout might be the perception of other shareholders that mutual funds are more likely to vote the same way as a block. Thus, other shareholders might perceive themselves to have a smaller impact on the election outcome and hence rationally decide to not take part in the elections. Literature attributes rational apathy to a lack of sufficient stake, a lack of ability to make an informed decision, and to the dispersion of ownership (Jill E. Fisch, 2017; Nili & Kastiel, 2016).

Following the literature, I define portfolio similarity measure as a dot product of corresponding portfolio vectors for a pair of mutual funds (Getmansky et al., 2018; Sias et al., 2013; Bohlin & Rosvall, 2014). This measure is also known as cosine similarity. I extend this measure to group similarity measure with a help of a two-step procedure. In the first

[9] Horizontal shareholding and diversification do not increase portfolio similarity measure if two investors diversify/hold competitors in two non-overlapping sets of companies. Since investors choose their holdings from the same universe of companies, these sets almost always overlap.

step, using mutual funds' shares in the company, I compute the weighted average portfolio of a group of mutual funds. In the second step, I compute a weighted average of similarity measures between this average portfolio and mutual funds' portfolio vectors. Thus, groups with closely related portfolios receive larger similarity measure values than groups with highly heterogeneous portfolios.

I study the effects of portfolio similarity at two different scales. First, I investigate at the level of individual funds by observing the relationship between portfolio similarity of a pair of mutual funds and probability that the pair casts the same votes. For every director election,[10] I select a small number of fund pairs at random from the pool of participating mutual funds.[11] This is a necessary data reduction step as considering every possible pair combination is not feasible for computation. I find a positive and statistically significant relationship.

Second, I investigate at the level of all mutual funds present at director elections. Using the portfolio similarity measure for groups, I find that mutual funds groups with more homogeneous portfolios cause the share[12] of votes "For" to drop and the share of "Non-votes" to rise.

Mutual funds with overlapping portfolios might be involved in self-selection into companies with more passive retail shareholders. At the same time, some retail shareholders might seek companies with more homogeneous institutional investor portfolios, so they can free ride on mutual funds' efforts in supporting good directors. Thus, I believe that my OLS results for the effect of mutual funds portfolio similarity on election participation might be biased.

[10] I concentrate on the sample of director elections because it is relatively homogeneous and abundant: most firms hold annual director elections with a number of positions to fill in.

[11] The possible bias from not choosing the non-participating funds should be small as mutual funds exercise their fiduciary duty by voting and the participation rate for institutional investors is very high (Jill E. Fisch, 2017; Nili & Kastiel, 2016).

[12] To be able to see the redistribution of shareholders' votes, I measure the share of particular voting option out of the total shares outstanding and eligible to vote during the meeting.

I use reconstitution of Russell 1000/2000 indices (FTSE Russell, 2019) to establish causality for the effect of mutual funds portfolio similarity on shareholders' voting decisions. Using instrumental variable approach, I attempt to capture exogenous variation in the degree of within-group portfolio similarity and the level of passive ownership. I instrument both variables in order to disentangle the effects of portfolio structure from the effects of ownership by index funds.

Since 2007, annual reconstitution of Russell 1000/2000 indices involves a banding procedure. Firms with market capitalization within a vicinity of the 1000th largest firm's market capitalization, do not switch the index.[13] I exploit both the inclusion of a firm in a certain index and its banded status to construct my instrumental variables. I use inclusion in Russell 2000 dummy, its lagged version, banded state dummy, and an interaction between banded state and inclusion in Russell 2000 as instrumental variables.

The weak instrument hypothesis is rejected using a test proposed by Sanderson & Windmeijer (2016). This test is specifically designed for the case of multiple endogenous variables. Conditional F-test statistics substantially exceed the 5% significance level's critical values for both endogenous variables.

The results of the instrumental variable approach largely confirm the main result of the OLS approach: more homogeneous groups of mutual funds, holding shares at a firm, reduce shareholder participation in director elections.

3.2 Data

The paper relies on three main kinds of data: portfolios of mutual funds, their voting decisions at corporate elections, and aggregate results of these elections. Additionally, I use data on characteristics of mutual funds and companies, and data on Russell 1000/2000 indices.

[13] The 1000th largest firm would be the threshold firm if sharp selection rule was used. The bandwidth is 5% of cumulative market capitalization of Russell 3000E.

As datasets come from a multitude of sources, I employ non-trivial automatic and manual matching that I describe in detail in this section and later in the Appendix.

My primary data source of mutual funds' voting decisions is the ISS's Voting Analytics database. This database provides individual mutual fund's votes that come from from N-PX filings, available on the EDGAR website. Since 2003, mutual funds have been required to publicly disclose their votes on all shares they hold. Along with mutual fund's vote, the Voting Analytics database attributes each mutual fund to a family of mutual funds and, starting in 2006, provides a link to the N-PX file that voting data was sourced from. The Voting Analytics database does not link its proprietary fund identifiers to these in other datasets, thus the links to underlying N-PX filings are very helpful in connecting the datasets.

The ISS's Voting Analytics database also contains aggregate voting data for Russell 3000 Index companies. Along with the company's CUSIP, numbers of votes "For," "Abstain," "Against/Withhold," and "Broker Non-votes," it also provides proposal's description, shareholder meeting date, management and ISS recommendations, sponsor information, number of shares outstanding, and the Pass/Fail outcome. Each proposal has a unique ID that allows to connect it to the votes of individual mutual funds.

The mutual funds' characteristics and portfolio data come from the CRSP Mutual Funds database. The portfolio composition data has quarterly frequency. For a mutual fund, which has voted at a corporate election, I use the holdings report that is nearest in terms of the absolute difference between the report date and shareholder meeting date (but no more than 183 days apart).[14] To ensure that result is not driven by artifacts of the CRSP MFDB dataset, I repeat the study with data from Thomson Reuters S12 dataset and get very similar results. Firm's characteristics are obtained from Compustat quarterly dataset.

[14]In some cases, for a mutual fund, which participated in a firm's shareholder meeting, there is no information on its share in the firm available in the selected CRSP MFDB's holdings report. In these cases I use adjacent holdings report for this fund to retrieve its approximate share in the firm.

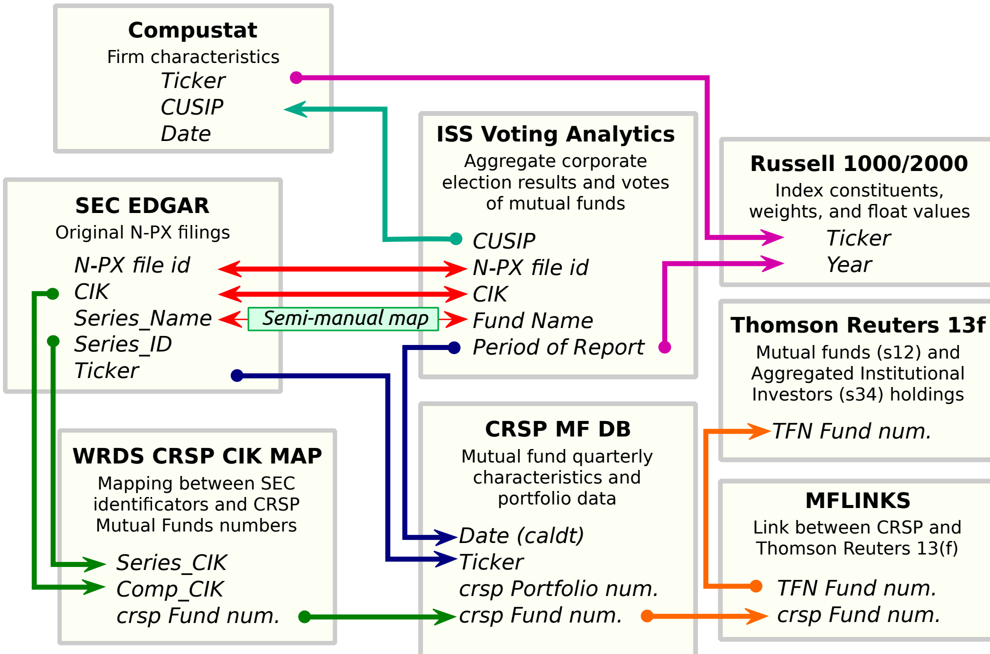

Figure 3.1: The datasets and the links between them. ISS Voting Analytics dataset provides aggregate corporate election results (numbers of votes "For," "Against/Abstain," "Withhold," "Broker Non-votes") along with short proposal description, ISS and management recommendations, and other vote-related information. I link it to Compustat by *CUSIP* and election date, and to Russell indices' constituents by the period of report and company's ticker, retrieved from Compustat. Linking ISS Voting Analytics dataset to CRSP Mutual Fund database of funds' characteristics and portfolios is done in two steps. First, I retrieve the original SEC N-PX filings from EDGAR web database for every mutual fund's vote in the ISS dataset. Since a N-PX form may contain voting data for more than one mutual fund, I match *Series_Name* to *Fund Name* from the ISS dataset. This allows me to associate a mutual fund from ISS dataset to its *Series_ID* and Ticker from SEC data. Second, I use mutual fund's ticker and date of the N-PX report to link the fund to its records in CRSP Mutual Fund database. There is also an alternative route that uses WRDS CRSP CIK MAP dataset that links pairs of *CIK* and *Series_ID (Comp_CIK)* to fund's records in CRSP Mutual Fund database. Both paths give very similar match results. Lastly, I match mutual fund records from CRSP Mutual Fund database to Thomson Reuters 13f (s12) database using MFLINKS dataset.

Russell 1000 and 2000 indices constituents, their "free float" share numbers, and the relevant stock prices are obtained from Bloomberg. I then compute the index weights and impute the ranks of index constituents.

Figure 3.1 presents a diagram of the links that I use to connect the datasets together. The most complicated step was to connect ISS Voting Analytics mutual funds voting dataset

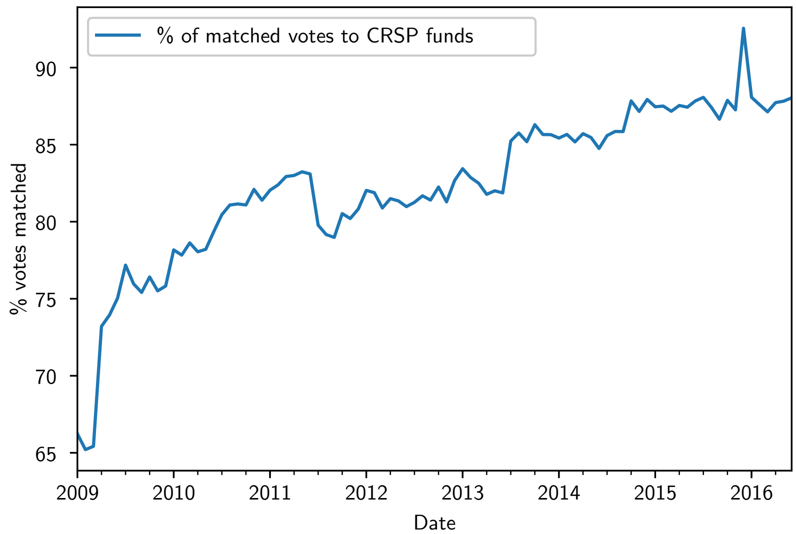

Figure 3.2: The average (monthly) match rate between the votes of mutual funds at director elections obtained from ISS Voting Analytics database and the mutual funds' portfolio identifiers from CRSP Mutual Fund database. For every director election in the ISS database, I retrieve all mutual funds that held shares of the respective company at the time of the election and reported these holdings in their N-PX filings. Then, for each such mutual fund, I attempt to find *crsp_fundno*, an identificator of respective mutual fund in the CRSP MF database, and *crsp_portno*, a portfolio identificator of the respective mutual fund in CRSP MF database. The matching procedure is described in the data section and in the Appendix. To find the match rate for a single director election, I divide the number of funds for which I was able to find corresponding portfolio identificator in CRSP MF database, *crsp_portno*, by the number of funds present in the ISS dataset for this director election. Then, I compute an average match rate for all director elections that happened within a calendar month and plot the figure above.

to CRSP Mutual Funds database. I follow the procedure outlined in Schwartz-Ziv & Wermers (2019); Matvos & Ostrovsky (2008) and Iliev & Lowry (2015). Figure 3.2 shows the match rate between the votes of mutual funds at director elections obtained from ISS Voting Analytics database and the mutual funds' records from CRSP Mutual Fund database. For years 2009 to 2016 I get a match of around 80% which motivates my choice of the time interval. I present the details of the match procedure in the Appendix.

3.3 Similarity measure

To compute the portfolio similarity measure for a pair of mutual funds I take a dot product of normalized vectors that represent the respective portfolios. In the literature this measure is known as cosine similarity. For groups of mutual funds I first compute a weighted-average portfolio vector and then evaluate the weighted-average similarity measure between the average portfolio and individual funds' portfolios.

To formally define the measure consider two mutual funds, i and j. Let vectors β_i and β_j represent the respective portfolio allocations. Then the portfolio similarity measure is

$$\text{Pair Similarity} = \mathbb{S}(\beta_i, \beta_j) = \frac{\langle \beta_i, \beta_j \rangle}{||\beta_i|| ||\beta_j||}, \tag{3.1}$$

where $|| \cdot ||$ is a L2-norm.

This measure has a clear geometrical representation. Consider a unit-sphere in a N-dimensional space, where N is the number of assets. Any portfolio, less of its scale, has a corresponding dot on the sphere. The normalized portfolio vector points from the origin to the point on the sphere. The dot product of a pair of such vectors is equal to the cosine of the angle between them. Smaller angles correspond to larger cosine values and portfolios' points being in a small vicinity of each other.

Theoretically, this measure spans from -1 (funds with completely opposite portfolios) to $+1$ (funds with identical up-to-scale portfolios), while in practice the range is $[0, 1]$ as short positions are not observed. Figure 3.3 illustrates the distribution of the measure values for a pair of mutual funds chosen at random. While most pairs have modest values, the distribution has a heavy tail and few pairs have values close to 1.

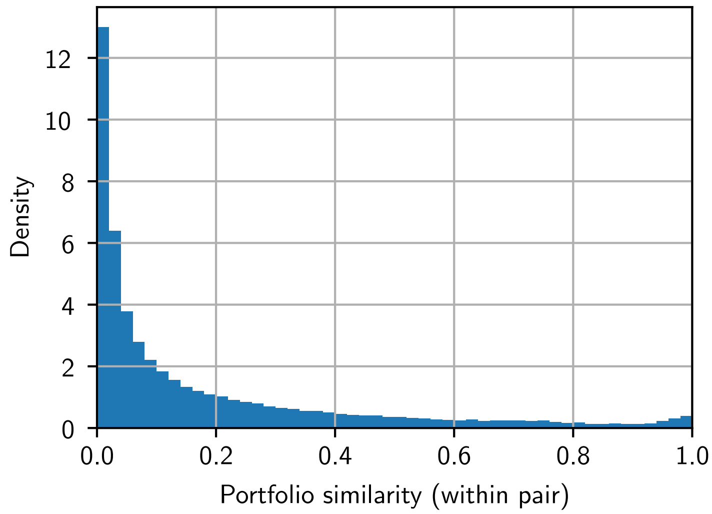

Figure 3.3: Distribution of portfolio similarity measure's (cosine similarity) values computed for 459507 random pairs of mutual funds at corporate elections from 2009 to 2016. For every director election between 2009 and 2016, I sample 3 random pairs (without return) of mutual funds. For every fund in a pair I retrieve portfolio data from CRSP MF database if I am able to match that fund to the respective portfolio. All pairs where one or more mutual funds miss portfolio data are discarded. Then, I compute the cosine similarity measure between the portfolios of funds in a pair.

Cosine similarity measure is widely known in the literature. The measure is used in portfolio analysis (Getmansky et al., 2018; Sias et al., 2013; Bohlin & Rosvall, 2014) and in text similarity analysis (Hanley & Hoberg, 2010, 2012).[15]

3.3.1 Similarity in groups

Cosine similarity measure accommodates the case of two mutual funds but it does not cover the case of many. I adapt the measure by computing a weighted average similarity between the weighted mean group portfolio and funds' portfolios.

The goal is to measure how diverse the group's portfolios are. When most shareholders have analogous portfolios, the mean portfolio will not be far from these. On the contrary, shareholders with heterogeneous portfolios will form a mean portfolio that is quite unlike

[15]Cha (2007) provides a survey of similarity measures. See Kwon & Lee (2003) and Sebastiani (2002) on usage of cosine similarity in text classification problems.

theirs. By measuring how similar their portfolios to the mean, I get an idea of how homogeneous the shareholders' portfolios are in the group.

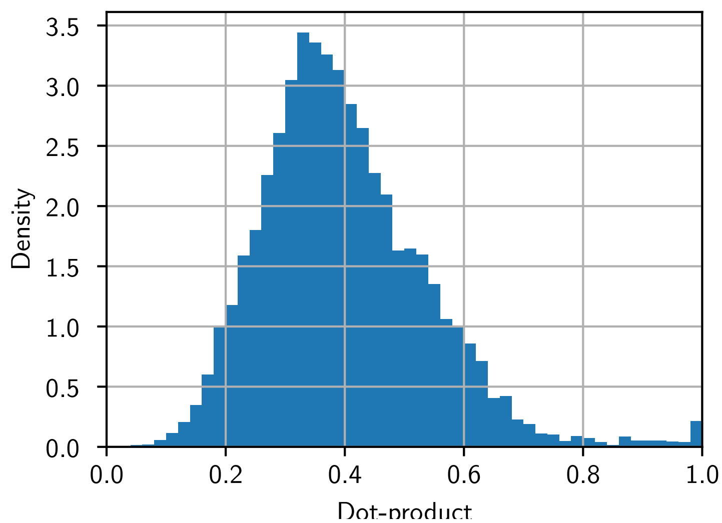

Figure 3.4: Distribution of group portfolio similarity measure's (group cosine similarity) values computed for mutual funds at corporate elections from 2009 to 2016. For every corporate election, I form a group of mutual funds which held shares of the respective company at the time of the election and reported these holdings in their N-PX filings. I match these mutual funds to their respective portfolios in the CRSP MF database. I use the mutual fund's share of the company normalized by the share owned by all mutual funds as the mutual fund's weight in the group similarity measure. In cases where the fund's most recent portfolio does not include shares owned in the company of interest, I use preceding or succeeding quarters of portfolio data to proxy for fund's holdings. Funds with missing portfolio data and funds with missing data on their holdings in the firm of interest are excluded from the group. Then, I compute the portfolio similarity measure among the funds remaining in the group.

I use shareholders' holdings at the company as the weights in this procedure. This ensures that many small shareholders of the company will not change the averages too much. To find the weighted mean portfolio of the group I weigh the shareholders' portfolios allocations (portfolio vectors with L1-norm being 1) by their shares at the company. Hence, the absolute size of shareholder's portfolio does not matter. Next, I use the same weights to compute the weighted average similarity measure for a group.

Consider a group of N shareholders, enumerated by i, with portfolio vectors β_i, where each element represents the dollar value allocated to shares of the respective company. Let $\mathbb{A}$ be the weighted mean group portfolio vector at the company n

$$\mathbb{A} = \frac{\sum_{i=1}^{N} \beta_{in} * \bar{\beta}_i}{\sum_{i=1}^{N} \beta_{in}}, \tag{3.2}$$

where $\bar{\beta}_i = \frac{\beta_i}{||\beta_i||_1}$ and $||\cdot||_1$ is L1-norm. Then the group similarity measure is

$$\text{Group Similarity} = \frac{\sum_{i=1}^{N} \beta_{in} \langle \mathbb{A}, \beta_i \rangle}{\sum_{i=1}^{N} \beta_{in}}, \tag{3.3}$$

where $\mathbb{S}$ is pair similarity measure defined above.[16]

Figure 3.4 illustrates how often we find a company where mutual funds, who are shareholders in that company, have certain level of portfolio similarity. Overall, this is a unimodal distribution with some outliers at the right tail.[17]

3.4 Similar portfolios and voting decisions

The voting outcome at corporate elections arises from decisions of many participants. While each of them might be too small to significantly affect the result, multiple participants following the same voting strategies may sway the election outcome. In this section I document that, among other factors, higher portfolio similarity in pairs of mutual funds is associated with greater probability of both funds making the same voting decision. This suggests that portfolio structure is likely related to the voting strategies of mutual funds in particular, and institutional investors in general.

[16] Another way to define group similarity measure could be the weighted average of pair similarity measure for all investor pairs. This would change the weighted mean group portfolio vector computation. That is, L1-norm will be replaced by L2-norm with all the rest being the same. Advantage of this approach is a more straightforward generalization from the 2 investors case, while the disadvantage is a more complicated definition of the group mean portfolio vector. In results I use the first approach. I've also implemented the second approach with results being very close to what the first approach yields.

[17] The value of zero is unattainable in the absence of short positions. The firms in the left tail of the distribution have mutual fund shareholders with almost non-overlapping portfolios.

To investigate what may affect shareholders' voting decisions I focus on votes of randomly chosen pairs of mutual funds that hold shares and vote at corporate elections.[18] For each fund in a pair, I collect fund characteristics, fund's share in a firm, and portfolio data. As a pair of funds does not have any order, both funds are equal participants in a pair. Thus, I do not directly use funds' characteristics and instead I construct averages and absolute differences. Using funds' portfolio data I compute the value of portfolio similarity measure. For each pair I also retrieve firm's characteristics.

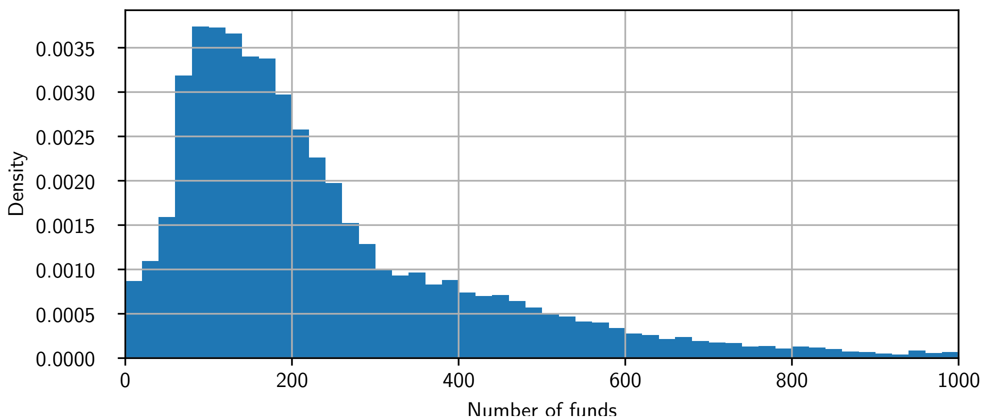

Figure 3.5: Distribution of the number of mutual funds participating in corporate elections from 2009 to 2016. Since I study all possible actions (including "Non-votes") a fund can take, I define participation as holding company's shares and reporting its vote (or "Non-vote") in an N-PX form by the mutual fund.

Shareholders may choose to vote "for," "against," "abstain," and "withhold." I consider two views on votes being the same: exact match and aggregation of "against," "abstain," and "withhold" votes. In the first approach, both funds in a pair must submit identical vote to count these as being the same. The second approach relaxes this condition a little bit: votes are considered the same as long as both funds submit votes from a single set, where sets are

[18]Typically, the number of mutual funds participating in a corporate election is relatively large: from 50 to 500 funds (see fig. 3.5). Thus, the number of possible funds pairs is in the range of hundreds to hundreds of thousands. Since I combine data from multiple election issues and multiple elections, the approach of looking at every possible pair quickly becomes infeasible. Instead, I randomly select mutual funds into a pair without return which drastically reduces the number of pairs that could be drawn. As this adversely affects the randomness of further pairs drawn, I only draw up to three pairs per election issue and in most cases I draw just one.

$\{\text{For}\}$ and $\{\text{Against, Abstain, Withhold}\}$. To test whether the second approach is reasonable, I also consider a placebo definition where I use sets $\{\text{Against}\}$ and $\{\text{For, Abstain, Withhold}\}$.

Corporate elections bring up a wide range of agendas: director elections, auditors ratification, compensation matters and say on pay votes, and proposals on governance matters. I concentrate my attention on director elections. The benefits involve the sample being quite homogeneous and abundant: almost every firm holds annual directors elections with multiple positions to fill in. The overwhelming majority of these directors run unopposed and are management-proposed.

For director elections from 2010 to 2016 I draw a single random pair of participating mutual funds. I use logistic regression to explain the relationship between the binary dependent variable, pair voted the same way, and the explanatory variables. I partially follow Iliev & Lowry (2015) in my selection of explanatory variables.

Table 3.1 presents the logistic regression results. Both, strict *(1)* and permissive *(2)* specifications yield statistically significant positive coefficient. For a logistic regression interpretation comes in a form of an odds ratio: the ratio of probability that funds make the same voting decision over the probability of making different decisions is about 41% higher with identical portfolios in comparison to a case of non-overlapping portfolios keeping other things fixed.

A pair of funds coming from the same mutual fund family is enormously more likely to vote the same way than a pair of funds from different families. This is consistent with literature that provides evidence of centralized voting behavior among the Big Three[19] mutual funds families (Fichtner et al., 2017).

The second most prominent effect comes from the ISS recommendation. A "For" recommendation from the ISS is associated with significantly higher chances of voting the same way. This is coherent with the literature on the role of the ISS and correlation between its recommendations and shareholders' votes (Choi et al., 2010; Iliev & Lowry, 2015; Malenko

[19]Big Three are BlackRock, Vanguard and State Street mutual fund families.

Table 3.1: Logistic regression: relationship between making the same voting decision and portfolio similarity for a randomly chosen pair of mutual funds. The dependent variable is dummy equal to 1 when votes are the same, and 0 otherwise. First specification requires votes to be exactly the same. Second relaxes the first by treating votes {Against, Abstain, Withhold} as being the same. Third specification is a placebo that treats votes {For, Abstain, Withhold} as being the same.

	Same vote (strict)	Same vote	For/Abs/Wth as a group
	(1)	*(2)*	*(3)*
Similarity measure (pair)	0.349***	0.348***	0.064
	(0.053)	(0.053)	(0.104)
Same family	3.007***	3.084***	3.004***
	(0.219)	(0.225)	(0.564)
1 index fund	0.057	0.056	−0.065
	(0.036)	(0.036)	(0.073)
2 index funds	0.059	0.053	−0.214**
	(0.050)	(0.050)	(0.103)
Same MSA	−0.092	−0.098*	−0.196*
	(0.059)	(0.059)	(0.109)
Geometric Averages of Funds Characteristics			
Expense ratio (geom. av.)	0.528***	0.531***	−0.067
	(0.076)	(0.076)	(0.142)
Management fee (geom. av.)	0.125	0.123	0.044
	(0.089)	(0.089)	(0.180)
Fund turnover ratio (geom. av.)	0.016	0.016	0.205***
	(0.037)	(0.038)	(0.080)
Total net assets (log(geom. av.))	−0.023**	−0.023*	−0.004
	(0.012)	(0.012)	(0.022)
Family size (log(geom. av.))	0.224***	0.223***	0.115***
	(0.011)	(0.011)	(0.020)
% of Total net assets (geom. av.)	0.025	0.029	0.139**
	(0.041)	(0.041)	(0.060)
% of Total Equity (geom. av.)	0.002	0.002	0.049
	(0.017)	(0.017)	(0.051)
Ratio of expense ratios	0.013***	0.013***	0.001
	(0.003)	(0.003)	(0.006)

Table 3.1, continued

	(1)	*(2)*	*(3)*
Absolute Differences of Funds Characteristics			
Management fee (abs. diff.)	0.170***	0.172***	0.002
	(0.038)	(0.038)	(0.069)
Fund turnover ratio (abs. diff.)	0.027**	0.026**	0.063**
	(0.011)	(0.011)	(0.026)
Total net assets (log(abs. diff.))	0.018**	0.019**	0.016
	(0.009)	(0.009)	(0.017)
Family size (log(abs. diff.))	−0.041***	−0.043***	−0.023
	(0.009)	(0.009)	(0.017)
% of Total net assets (abs. diff.)	−0.030**	−0.031**	−0.036
	(0.015)	(0.015)	(0.024)
% of Total Equity (abs. diff.)	0.012**	0.012**	−0.008**
	(0.005)	(0.005)	(0.003)
Firm characteristics and ISS Recommendations			
S&P 500	0.115**	0.110**	−0.496***
	(0.047)	(0.047)	(0.089)
ISS Against another item	−0.377***	−0.380***	−0.106*
	(0.030)	(0.030)	(0.062)
ISS "For" recommendation	2.493***	2.474***	1.884***
	(0.034)	(0.034)	(0.074)
log(Total assets)	0.102***	0.103***	−0.221***
	(0.011)	(0.011)	(0.022)
Return on assets	−0.001	−0.001	−0.008*
	(0.001)	(0.001)	(0.004)
Book to market ratio	−0.116***	−0.111***	0.041
	(0.023)	(0.023)	(0.068)
Leverage	−0.018	−0.019	0.047*
	(0.012)	(0.012)	(0.025)
Constant	−2.799***	−2.770***	3.352***
	(0.167)	(0.167)	(0.320)
Observations	85099	85099	85099

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

& Shen, 2016). Notably, an ISS recommendation against another item is linked to reduced chances of vote unanimity within the mutual funds pair.

Funds with higher expense ratios and from bigger mutual fund families tend to cast similar votes. Firms in S&P 500 and larger firms in general receive a more homogeneous treatment from mutual funds, whereas firms with higher book to market ratio are more likely to receive different votes.

The placebo specification which combines "For," "Abstain," and "Withhold" votes performs largely as expected. Portfolio similarity has no significant association with unanimity in pair's votes. Same family and ISS recommendation dummies retain relatively large albeit less significant coefficients. This is likely due to their resolutive power between "For" and "Against" votes even when "Abstained" and "Withhold" votes somewhat scramble the pair's outcomes.

3.5 Similar portfolios and shareholder participation

As individual pairs of mutual funds more likely to cast same votes having analogous portfolios, the same effect should scale to groups of many mutual funds. With higher probability of them voting the same way, other shareholders may re-evaluate their decision to participate in elections. I find that at elections where mutual funds, as a group, have more similar portfolios, other shareholders decide to not submit their votes. This shrinks the pool of votes cast, which in turn may enhance the power of those who vote.

To see the relationship between portfolio similarity and shareholder participation I investigate how portfolio similarity is related to aggregate characteristics of elections' outcomes. As in the previous section, I focus on director elections to benefit from abundance and homogeneity of these as well as to maintain consistency within the paper.

3.5.1 Voting standards

The corporate governance process through shareholder voting is covered with a patchwork regulatory framework composed of federal and state corporate and securities laws, stock exchange requirements and company bylaws. Director elections are usually governed by a state law default if company's bylaw provides no other standard (Stokdyk & Trotter, 2016). The two most used standards are majority voting and plurality voting.

Shareholders have a variety of options at director elections. They can support the candidate by voting "For," disapprove the candidate by voting "Against" (or "Withhold" under plurality voting standard), or be more neutral and vote "Abstain." Another option is to do nothing and do not vote at all.[20] This would result in a "Non-Vote," an outcome that covers the case not covered by the options above. Depending on the voting standard, the voting options have different effect on the election outcome.

Under a plurality voting standard, director candidate who receives the highest number of votes "For" wins the seat. Notably, if candidate is running unopposed, a single vote "For" is enough to get elected. Shareholders may wish to vote "Withhold" if they are not happy with the candidate. While high number of withhold votes does not prevent such candidate from being elected, the board of director may adjust its director nomination practices (The Office of Investor Education and Advocacy, 2012).

Under a majority voting standard, director nominee needs to secure enough "For" votes to satisfy a majority voting requirement. The requirement typically describes a threshold that the share of "For" votes needs to pass. Table 3.2 shows the subsets of votes used to calculate the share of "For" votes under different standards. These serve as denominator in a formula used to compute the support rate. The ISS Voting Analytics dataset suggests that less than 1% of director elections use majority of outstanding shares as base. For example, GMI Ratings (2013) report that 94% of companies in both S&P 500 and Russell 1000 exclude broker non-votes for shareholder proposals.

Majority voting standard does not preclude unpopular directors from getting elected. This standard has been on the rise since 2004 and by 2007 approximately two-thirds of S&P

[20] "Abstain" vote is an affirmative choice of a shareholder, represented at the meeting (by proxy or in person), to not vote "For" or "Against" particular candidate. Abstentions may or may not be considered "vote cast." See Schnell & Chen (2019).

Table 3.2: Types of votes forming the base that is used to compute the share of votes "For" under different majority voting standards.

Standard	For	Against	Abstain	Broker Non-Vote	Other Non-Vote
Majority of votes cast	✓	✓			
Majority of shares present and entitled to vote on the subject matter (Default in Delaware)	✓	✓	✓		
Majority of shares present and entitled to vote at the meeting	✓	✓	✓	✓	
Majority of outstanding shares	✓	✓	✓	✓	✓

500 firms used some form of majority voting (Allen, 2007). The adoption was not uniform; at the early stages firms were also introducing a "plurality plus" standard which required elected candidate to resign, pending an approval of resignation from board of directors, if he or she fails to win under a majority vote standard. Cai et al. (2013) claims that majority voting standard is a paper tiger, instituted to appease shareholders, which has little teeth to affect director elections. Cai et al. (2009) finds that even at poorly performing firms with bad governance badly performing directors consistently receive more than 90% of votes "For"; the exceptions are negative ISS recommendation with 19% fewer votes and directors attending less than 75% of board meetings who receive 14% fewer votes. Hence, in virtually every case even an unpopular director receives more than 50% of votes "For" which brings him above the usual threshold in majority voting elections. Cai et al. (2013) find that from 2004 to 2010 from 105445 directors only 294 directors at 153 firms received less than 50% of "For" votes; among these 153 firms only 14 firms had adopted a version of majority voting standard and all except 3 of 22 directors that failed elections at these firms secured a seat

on the board. Thus, majority voting standard is not much different from plurality voting standard in the case of director elections.

While elections do not weed out unpopular candidates, dissent votes convey a credible signal of shareholder displeasure to board of directors and management. Yermack (2010) argues that this signal may pressure management to change the composition of the board, dismantle takeover defenses, and to revise executive compensation packages. Cai et al. (2009) estimate that a 1% decrease in support of compensation committee member tends to reduce unexplained CEO compensation by $143,000 in the following year. They also find that CEO turnover is more likely when independent directors receive lower votes. Iliev et al. (2015) show an association between low percentages of "For" votes and a higher number of directors leaving the board over the next year. So even if director elections have no immediate effect, there is an evidence of delayed action taken by the board and firm management.

Thus I do not make a distinction between the two voting standards as in both cases shareholders have a way to display their dissent and the majority voting standard poses no substantial barriers for directors to get on the board. At the same time election results matter for the corporate governance in the long run: low percentages of "For" votes nudge management and the board into taking shareholder appeasing actions.

3.5.2 Data sample

The ISS Voting Analytics dataset provides vote results for corporate elections that include vote outcome, number of outstanding shares, number of votes "For," "Abstain," "Against"/"Withhold,"[21] and number of broker non-votes[22] along with individual votes of mutual funds and other election information.

[21]ISS does not provide a separate number of "Withhold" votes for records starting 2006; instead it reports this number in the "Against" column.

[22]A broker non-vote happens when a beneficial owner does not submit voting instructions to a broker through which she holds shares. Brokers, in general, are permitted to vote on behalf of beneficial owners on "routine" matters without explicit instructions from beneficial shareholders but director elections are not considered "routine."

Table 3.3: Director elections sample's summary statistics. Sample contains 95543 election datapoints after removing outliers and records with missing values.

Variable	Min	Max	Mean	Median	std.
"For" votes, %	3.000	99.782	74.694	77.022	12.810
"Against"/"Withhold" votes, %	0.000	71.375	3.330	1.370	5.667
"Abstain" votes, %	0.000	39.112	0.146	0.000	0.785
Broker "Non-Votes", %	0.000	69.596	9.733	7.953	7.715
Other "Non-Votes", %	0.000	96.989	12.098	10.657	8.347
Similarity measure	0.151	0.794	0.397	0.375	0.113
log(Total assets)	2.615	14.637	7.973	7.909	1.942
Return on assets, %	−49.985	49.844	3.122	3.134	9.264
Book to market ratio	0.001	16.544	0.567	0.479	0.477
Leverage	0.000	9.880	0.870	0.516	1.200
% owned by index funds	0.000	68.362	7.872	7.519	4.302
% owned by non-index funds	0.000	95.818	16.274	15.907	9.479
ISS "For" recommendation	0.000	1.000	0.923	1.000	0.266

For every director election happened between 2009 and 2016 I collect its aggregate voting outcomes together with the individual votes of mutual funds from ISS Voting Analytics dataset. Using the numbers of votes cast I construct the number of "Non-votes" which is the difference between "Shares Outstanding" and total number of votes cast. The number of "Non-votes" is then split into "Broker non-votes," reported by the ISS, and "Other non-votes." I drop election issues where number of votes cast exceeds the reported number of shares outstanding.

Using CRSP Mutual Funds database and Thomson Reuters S12 database I retrieve mutual fund portfolios as well as their shares in the firms that they vote at.[23] Next, I evaluate the group portfolio similarity measure for mutual funds participating in director elections. The firm characteristics are then pulled from Compustat. Finally, I aggregate votes of mutual funds to evaluate their input into the aggregate voting election results. Table 3.3 provides summary statistics for the constructed sample.

[23]The portfolio data is updated quarterly so for some mutual funds I can not directly observe its holdings at the firm of interest. In such cases I use adjacent quarters data to proxy for the missing number.

3.5.3 Results

Mutual funds portfolio similarity has a sizable association with directors election outcomes. Firms where mutual funds, as shareholders, have more similar portfolios are more likely to see lower shareholder participation in director elections. Fewer votes "Against"/"Withhold" and substantially fewer "For" votes being cast, while shareholders increase the share of "Non-Votes" by not submitting their voting instructions.

To explore the relationship between portfolio similarity and election outcomes I regress the numbers of votes "For," "Against," "Abstain," and "Non-votes," normalized by the shares outstanding, on similarity measure, shares of index and non-index mutual funds, ISS recommendation, and firm controls. The normalization used allows me to see how shareholders dispose their votes across all possible options. Table 3.4 presents the OLS regression results.

The percentage of votes "For," measured as a share of shares outstanding, changes the most in relation to portfolio similarity. As these five dependent variables cover the entire set of possible outcomes[24], a drop in "For" votes should be accompanied with a hike in other categories. The share of "Against" votes decreases as well,[25] thus the missing "For" and "Against" votes end up as "Non-Votes" as shareholders reduce their involvement in the election process. As both "Broker Non-Votes" and "Other Non-Votes" have substantial positive coefficients for portfolio similarity, I can conclude that both retail and institutional investors' decisions are at play.

More homogeneous groups of mutual funds, that hold shares in a firm, thereby decrease[26] shareholder turnout. Jill E. Fisch (2017) argues that low turnout among retail shareholders leads institutional investors to dominate election results; she reports that while 90% of insti-

[24] ISS uses "Against" column to report "Withhold" votes when needed and I follow this practice here.

[25] While the coefficient is about an order of magnitude smaller, the effect on "Against" votes is still substantial as the average number of votes "Against" is also more than a magnitude smaller than the number of votes "For." See table 3.3 for a detailed summary statistics.

[26] Here I will talk about possible causal effects that a lower shareholder turnout may have. I leave the discussion of whether portfolio similarity causes shareholder turnout to drop for a separate section.

tutional shares are voted, retail investors turnout averages at less than 30% (see also Matt Egan (2014)). Nili & Kastiel (2016) claim that retail investors have rational apathy which stems from the dispersion of ownership and diversification of investor portfolios, and cite vote outcome distortion, limitation of shareholders' ability to initiate governance changes, and dead-lock situations where low shareholder turnout prevents issues from passing as the direct costs of investors' apathy.

Election outcome is also responsive to the share of a firm owned by mutual funds. Using the CRSP Mutual Funds database, I'm able to disentangle the input of index mutual funds[27] and non-index mutual funds. Higher share, owned by index mutual funds, reduces observed number of votes "For" and increases the number of votes "Against" as shareholders tend to not vote their shares. Quite the opposite happens when non-index mutual funds hold bigger share: shareholders, institutional and not, tend to vote more often.

ISS issues a favorable recommendation for director elections in more than 90% of the cases. A lack of favorable ISS recommendation has an expected relationship to the election outcome: a sharp decrease in the number of votes "For" coming from an increase of even bigger magnitude in number of votes "Against." Part of these votes "Against" come from previously passive shareholders as can be seen from a decrease in the numbers of non-votes.

Larger firms attract more retail shareholder attention. Higher return on assets is associated with better shareholder participation, while high leverage is associated with lower.

3.6 Instrumental strategy

Since mutual fund's portfolio choice is likely endogenous to its voting behavior, I instrument group portfolio similarity measure to establish causality. I use reconstitution of Russell 1000/2000 indices as a source of portfolio variation that is plausibly exogenous to share-

[27]I consider a mutual fund as an index fund if it has flag "D" in the `findex_fund_flag` field of CRSP Mutual Fund database.

Table 3.4: The relationship between mutual funds portfolio similarity and election outcome at director elections. Dependent variables normalized by the shares outstanding. Standard errors are robust to cluster correlation (clustered by meetings).

	% For	% Against	% Abstain	% Broker Non-Vote	% Other Non-Vote
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Similarity measure	-14.151^{***}	-2.448^{***}	0.124^{**}	7.218^{***}	14.892^{***}
	(1.086)	(0.359)	(0.063)	(0.674)	(0.713)
% owned by index funds	-0.489^{***}	0.149^{***}	0.001	0.265^{***}	0.127^{***}
	(0.032)	(0.011)	(0.002)	(0.020)	(0.020)
% owned by non-index funds	0.417^{***}	0.050^{***}	-0.000	-0.228^{***}	-0.213^{***}
	(0.012)	(0.004)	(0.001)	(0.008)	(0.007)
ISS “For” recommendation	10.404^{***}	-14.372^{***}	-0.083^{***}	2.159^{***}	3.668^{***}
	(0.405)	(0.250)	(0.032)	(0.174)	(0.215)
log(Total assets)	0.264^{***}	-0.018	0.028^{***}	-0.474^{***}	0.464^{***}
	(0.063)	(0.019)	(0.005)	(0.041)	(0.042)
Return on assets, %	0.194^{***}	-0.005	-0.001	-0.085^{***}	-0.113^{***}
	(0.012)	(0.004)	(0.001)	(0.008)	(0.008)
Book to market ratio	-0.007^{***}	0.001^{***}	-0.000^{*}	0.006^{***}	0.001^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Leverage	-0.185^{*}	-0.108^{***}	0.008	0.520^{***}	-0.238^{***}
	(0.106)	(0.026)	(0.011)	(0.076)	(0.062)
Year controls	Yes	Yes	Yes	Yes	Yes
Observations	95543	95543	95543	95543	95543
R^2	0.222	0.434	0.005	0.252	0.710
F stat.	240.3	248.7	19.4	861.9	3329.8

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

holders' voting practices. The instrumental approach confirms the findings of the previous section.

There are multiple reasons why the OLS results may not be conclusive evidence that mutual funds' portfolio structure affects election outcomes. One possibility is that the highly diversified mutual funds might be better represented at "boring" firms where shareholders do not often engage in director elections. This way low shareholder turnout might be correlated with pools of highly diversified, but essentially holding very similar portfolios, mutual funds.

Another possibility is that mutual funds endogenously determine their portfolios. Funds' ownership of a stock might be related to factors that directly affect shareholder turnout. This way a correlation between portfolio similarity and election participation might not represent a causal relation.

To address the possible endogeneity, I exploit reconstitution of Russell 1000 and Russell 2000 indices, widely adopted market benchmarks, as a quasi-natural experiment in changing the stock ownership by mutual funds. A substantial difference in the index-weight of a stock at the top of Russell 2000 and the bottom of Russell 1000, as well as probability of switching an index, drives the changes in portfolios of investors, which rely on Russell indices in their portfolio building.

3.6.1 Russell 1000/2000 indices reconstitution

Russell 1000 and Russell 2000 indices, provided by FTSE Russell, are stock market indices that track the highest-ranking 1000 stocks and stocks ranking 1001 - 3000 respectively. Many investment managers use Russell 2000 to benchmark their performance in "small-cap" to "mid-cap" categories and build portfolios. Every year in May - June, FTSE Russell reconstructs the indices, revealing the result on the last Friday in June.

Historically, there have been two different procedures for Russell 1000/2000 indices reconstitution. Prior to 2007, index assignment followed a strict threshold rule where stocks ranked within the 1-1000 interval were assigned to Russell 1000, and stocks ranked in 1001-

3000 were assigned to Russell 2000. Beginning in 2007, a new approach, called "banding," was enacted. The threshold between Russell 1000 and Russell 2000 is now covered with a band of a certain dollar size, such that companies that fall into the band at reconstitution do not switch the index. As I concentrate on a data sample from 2009 to 2016 in this paper, I will skip the discussion of former approach to indices reconstitution and I will focus on the latter.

Reconstitution starts by forming a ranked list of 3000+ companies that are eligible to be included in one or more of the Russell indexes. Few criteria for eligibility among others are having stock price above \$1, being a part of the U.S. equity market, having total market capitalization above \$30 million, and having more than 5% of shares available in the marketplace (float) (FTSE Russell, 2019). Total market capitalization of a firm is obtained by multiplying total outstanding shares by the market price (last price traded) on the primary exchange on the rank day[28] in May. FTSE Russell estimates the total shares outstanding by including common stock, partnership units/membership interests, and non-restricted exchangeable shares, while any other type of shares (preferred stock, installment receipts, etc.) are excluded (see FTSE Russell (2019) for a detailed description). Computed total market capitalization allows Russell to sort companies in a long 3000+ list where position in the list determines the rank of a company. FTSE Russell treats the computed market capitalizations as proprietary information and does not make it available to researchers. This hampers the research trying to exploit reconstitution of Russell indices in a regression discontinuity design setting (Wei & Young, 2017). The more recent papers (Heath et al., 2018) focus on the post 2006 period and provide new methodology, immune to selection bias.

Next, a band with a width of 5% of cumulative market cap of Russell 3000E is computed around the market capitalization of security ranked 1000. Any company which total market capitalization falls within the band does not switch the index. Companies outside the band and with ranks below 1000 become Russell 1000 constituents, while such companies with

[28]Schedule of rank days is released by Russell in spring and, typically, rank day is the last trading day in May (Mullins, 2014).

ranks above 1000 become Russell 2000 constituents. Thus, banding procedure provides an additional signal to investors regarding company's future affiliation with a certain index.

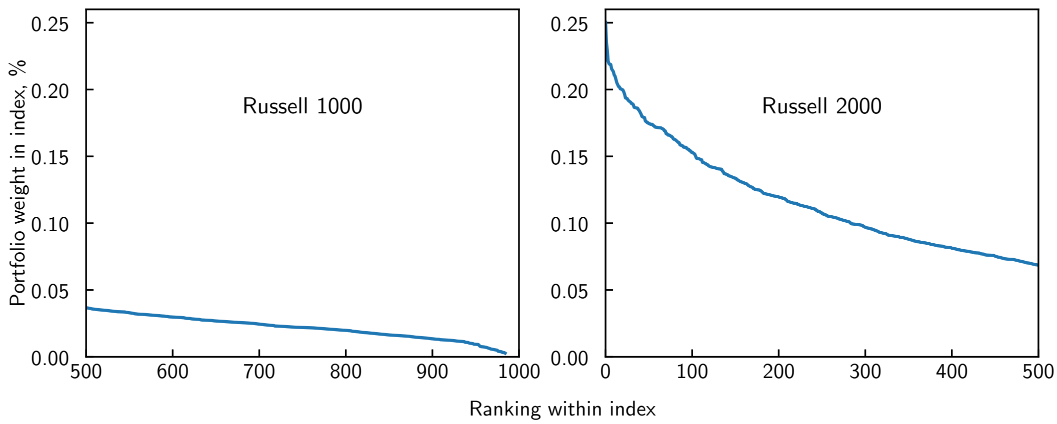

Figure 3.6: Portfolio weights of stocks in Russell 1000/2000 indices in 2011. Stock's weight within an index is computed by dividing the stock's "float market capitalization" by the sum of "float market capitalizations" of all index constituents. "Float market capitalization" is a product of the end-of-June stock's share price and the number of shares that can be freely traded by the public. Stock's rank within an index is equal to its position in a list of all index constituents, ordered by their weights within the index starting with the largest value.

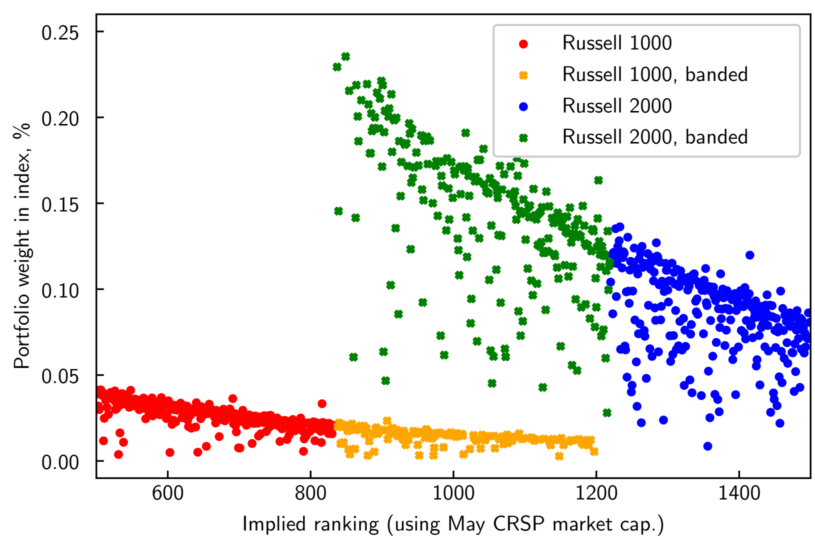

Figure 3.7: Portfolio weights of stocks in Russell 1000/2000 indices in 2011. FTSE Russell determines securities ranks by using proprietary data on firms' total market capitalizations on the rank day in May. I compute implied ranks by using May CRSP firms' market capitalization data.

Once the indices memberships have been determined, FTSE Russell adjusts companies' shares to only include those that can be freely traded by the public ("free float"). Next, within each index separately the float adjusted shares are used to compute constituents' weights. Given the sorting nature of the indices, companies at the bottom of Russell 1000 index receive substantially lower weights in comparison to the weights of companies at the top of Russell 2000 index. Figure 3.6 illustrates the difference between the weights in an index for the bottom 500 companies from Russell 1000 and the top 500 companies of Russell 2000 in 2011. The average weights of securities at the bottom and top 500 companies of respective indices are 0.023% and 0.118% respectively.

3.6.2 Exclusion restriction

An important criterion for instrument variable estimation approach is exclusion restriction. I construct four instrumental variables that are based on the features of indices reconstitution. In this section I explore the possible critiques of these instruments.

The first pair is a dummy that a certain security belongs to the Russell 2000 index and its lagged version. Since index reconstitution happens mid-year, for every election before index reconstitution date I use values from previous two years, while after reconstitution date I use the current and the past year's values.

The second pair is the banded status of a firm and an interaction term between the banded status and the inclusion in Russell 2000 index. The banded status is effectively a proxy for a possible index switch in a future. Wei & Young (2017) hypothesize that institutional investors may trade in anticipation of index assignment changes.

In an instrumental regression, I rely on an implicit assumption that these instrumental variables affect the dependent variables (votes cast at elections) only through its influence on the variables of interest. Literature suggests that passive and active funds have a different impact on firm's governance and corporate election outcomes (Appel et al., 2016, 2018; Schmidt & Fahlenbrach, 2017; Brav et al., 2019; Heath et al., 2018). Russell indices reconstitution has also been associated with a significant changes in index fund ownership at firms that switch the index (Appel et al., 2016; Heath et al., 2018; Gloßner, 2018). Thus, I control for levels of institutional ownership by including both passive and active ownership shares in my regression. Then, I go the extra mile by exploiting the different nature of the available instruments in order to instrument for both the portfolio similarity and the level of passive ownership. This allows me to disentangle the effect coming from portfolio similarity from the effects found in the literature studying passive ownership influence (Appel et al., 2018; Heath et al., 2018; Baig et al., 2018). Following Appel et al. (2018) I also control for market capitalization and free float of a firm.

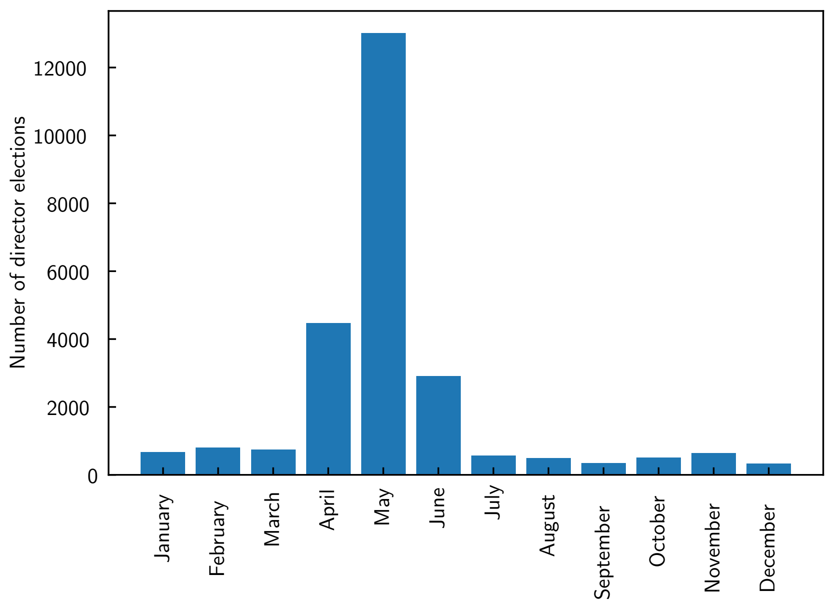

Figure 3.8: Number of director elections per month in the IV sub-sample.

One possible critique is that index switching may generate news coverage or in some other way attract (or reduce) shareholder attention to a company. Thus, some companies may enjoy less or more shareholder election participation just due to the switch itself. I do not find this concern substantial as the bulk of director elections happens many months after the index reconstitution and any news effect should be worn away by then. Figure 3.8

provides an illustration that most director elections are over by the end of June when the new lists of constituents become public.

Another possibility for a violation of exclusion restriction arises if firms are able to manipulate their index assignment. Then their actions might affect not only the index assignment but also the shareholders' attitude for election participation. As FTSE Russell does not reveal the true May rankings, firms have small chances to successfully predict their rank and then to manipulate the index they belong to or their banded status by a marginal change in capitalization. The articles that explore Russell reconstitution in regression discontinuity design setting do not find evidence of firms manipulating their index assignment (Boone & White, 2015; Chang et al., 2015).

Finally, since the number of instruments used makes the model overidentified, I perform a Sargan–Hansen test for regressions with statistically significant influence of portfolio similarity (Hayashi, 2000). In all cases the resulting statistic is sufficiently smaller than the one needed in order to reject the hypothesis of the over-identifying restrictions being valid. This strengthens my belief that the exclusion restriction is satisfied.

3.6.3 Weak instruments test

An ambitious goal to instrument two endogenous variables requires a careful attention to the strength of the instruments being used. A simple F-test will not be sufficient in such case as it does not capture the interplay between the endogenous variables. I rely on a test proposed by Sanderson & Windmeijer (2016) to rule out weak instruments case.

Table 3.5 summaries the results of the first stage. Both regressions demonstrate an F-statistic above 10. Thus, the instrument are strong enough at least in the case when just one variable is instrumented.

Table 3.5: The first stage of the 2SLS regression. Standard errors are robust to cluster correlation (clustered by meetings).

	Similarity measure (1)	% owned by index funds (2)
$Russell2000_t$	0.020**	1.466***
	(0.008)	(0.330)
$Russell2000_{t-1}$	0.010*	−0.273
	(0.006)	(0.223)
Banded state	−0.001	−0.303
	(0.005)	(0.202)
Banded state $\times Russell2000_t$	−0.029***	0.353
	(0.007)	(0.314)
Firm controls	Yes	Yes
Year controls	Yes	Yes
Float and mk.cap. controls	Yes	Yes
Observations	29167	29167
R^2	0.238	0.467
Partial F stat.	55.0	177.4

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Sanderson & Windmeijer (2016) argues that in a case of multiple endogenous variables being instrumented, a simple F-test is necessary but not sufficient. They provide a conditional F-test statistic which I compute for my first stage. The values are 79.8 and 94.8 for similarity measure and share of index fund ownership respectively. Stock, J.H. (2005) provide the 5% significance level's critical values for a weak instruments test. The null hypothesis is that instruments are weak and lead to an asymptotic bias of at least 5%. The critical value for four instruments and two endogenous variables is 11.04. Since both values are substantially higher than the critical value, the null hypothesis of weak instrument is rejected at the 5% level.

3.6.4 Results

Table 3.6 summarizes the results of IV approach. The main result on the influence of portfolio similarity remains in place. The substantial drop in the share of votes "For" combined with a hike in the number of broker "Non-Votes" suggest the similar pattern of shareholder fatigue and reduced participation in director elections. A notable sign change has happened for the effect coming from share owned by index funds; IV results suggest than an increase in this share raises retail shareholder participation and increases the level of support directors receive at elections.

Among other variables, the changes came from the logarithm of total assets and book to market ratio. The former variable lost its significance in the regression while the latter obtained much more pronounced effects likely due to inclusion of float and market capitalization control variables. The effect of favorable ISS recommendation largely remains the same.

In tables 3.7 and 3.8, I modify the group similarity measure by only including shares within the same 1 or 2 digit SIC category. This allows me to see the effect of horizontal shareholding on voting outcomes. Both regressions' results are in coherence with results in table 3.6.

3.7 Conclusion

According to my study of mutual funds' portfolios and voting patterns, portfolio structure has an effect on both individual voting behavior and aggregate outcome of director elections. Funds with more similar portfolios tend to cast identical votes more often. I find that greater within-group portfolio similarity of mutual funds, invested in a firm, causes lower shareholder participation in director elections at this firm.

The observed relation between portfolio structure and individual fund's voting behavior provides evidence that mutual funds exercise their own judgment to some extent and do not

Table 3.6: Relationship between mutual funds portfolio similarity and election outcome at director elections (IV approach). Dependent variables normalized by the shares outstanding. Standard errors are robust to cluster correlation (clustered by meetings).

	% For	% Against	% Abstain	% Broker Non-Vote	% Other Non-Vote
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Similarity measure	-40.857^{**}	6.845	-0.407	22.015^{**}	12.404
	(16.723)	(6.261)	(1.093)	(10.659)	(10.136)
% owned by index funds	0.675^{**}	-0.200^{*}	0.012	-0.442^{**}	-0.045
	(0.320)	(0.112)	(0.019)	(0.207)	(0.200)
% owned by non-index funds	0.285^{***}	0.052^{***}	-0.002	-0.148^{***}	-0.187^{***}
	(0.053)	(0.020)	(0.003)	(0.033)	(0.034)
ISS "For" recommendation	15.855^{***}	-17.422^{***}	-0.049	0.831^{***}	0.785^{**}
	(0.745)	(0.556)	(0.043)	(0.320)	(0.346)
log(Total assets)	-0.376	0.023	0.002	0.252	0.099
	(0.342)	(0.119)	(0.032)	(0.222)	(0.221)
Return on assets, %	0.076^{***}	0.022^{**}	0.001	-0.019	-0.080^{***}
	(0.029)	(0.011)	(0.002)	(0.020)	(0.020)
Book to market ratio	-2.593^{***}	0.286^{*}	0.159^{***}	0.963^{***}	1.185^{***}
	(0.530)	(0.167)	(0.061)	(0.338)	(0.387)
Leverage	-0.591^{***}	-0.123^{**}	0.022	0.748^{***}	-0.056
	(0.184)	(0.060)	(0.015)	(0.133)	(0.083)
Year controls	Yes	Yes	Yes	Yes	Yes
Float and mk.cap. controls	Yes	Yes	Yes	Yes	Yes
Observations	29167	29167	29167	29167	29167
R^2	0.231	0.462	0.011	0.189	0.136
F stat.	1250.2	1245.1	55.7	1545.3	573.4

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 3.7: Relationship between mutual funds portfolio similarity (using only assets within the same 1-digit SIC category) and election outcome at director elections (IV approach). Dependent variables normalized by the shares outstanding. Standard errors are robust to cluster correlation (clustered by meetings).

	% For	% Against	% Abstain	% Broker Non-Vote	% Other Non-Vote
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Similarity measure (1 digit SIC)	-56.051^{**}	8.930	-0.528	29.071^{**}	18.578
	(23.353)	(8.301)	(1.446)	(14.681)	(13.463)
% owned by index funds	0.901^{**}	-0.259^{**}	0.014	-0.535^{**}	-0.121
	(0.371)	(0.126)	(0.021)	(0.231)	(0.223)
% owned by non-index funds	0.319^{***}	0.049^{***}	-0.002	-0.170^{***}	-0.196^{***}
	(0.045)	(0.016)	(0.003)	(0.027)	(0.028)
ISS "For" recommendation	15.140^{***}	-17.190^{***}	-0.054	1.131^{***}	0.973^{***}
	(0.818)	(0.558)	(0.046)	(0.362)	(0.369)
log(Total assets)	-0.356	0.023	0.002	0.217	0.114
	(0.357)	(0.117)	(0.031)	(0.229)	(0.216)
Return on assets, %	0.102^{***}	0.015	0.001	-0.031^{*}	-0.087^{***}
	(0.027)	(0.009)	(0.001)	(0.017)	(0.018)
Book to market ratio	-2.529^{***}	0.272	0.159^{***}	0.933^{***}	1.166^{***}
	(0.549)	(0.171)	(0.061)	(0.350)	(0.384)
Leverage	-0.610^{***}	-0.119^{**}	0.022	0.759^{***}	-0.052
	(0.195)	(0.061)	(0.015)	(0.137)	(0.084)
Year controls	Yes	Yes	Yes	Yes	Yes
Float and mk.cap. controls	Yes	Yes	Yes	Yes	Yes
Observations	29233	29233	29233	29233	29233
R^2	0.125	0.446	0.008	0.099	0.113
F stat.	1080.8	1196.9	56.7	1396.8	581.5

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 3.8: Relationship between mutual funds portfolio similarity (using only assets within the same 2-digit SIC category) and election outcome at director elections (IV approach). Dependent variables normalized by the shares outstanding. Standard errors are robust to cluster correlation (clustered by meetings).

	% For	% Against	% Abstain	% Broker Non-Vote	% Other Non-Vote
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Similarity measure (2 digits SIC)	-65.216^{*}	7.391	-0.792	36.597^{*}	22.020
	(35.926)	(10.006)	(1.680)	(21.819)	(17.475)
% owned by index funds	1.210^{**}	-0.276^{*}	0.019	-0.725^{**}	-0.228
	(0.561)	(0.157)	(0.026)	(0.338)	(0.289)
% owned by non-index funds	0.358^{***}	0.041^{***}	-0.002	-0.189^{***}	-0.209^{***}
	(0.041)	(0.012)	(0.002)	(0.025)	(0.022)
ISS "For" recommendation	15.313^{***}	-17.230^{***}	-0.054	1.055^{**}	0.916^{**}
	(0.936)	(0.562)	(0.045)	(0.417)	(0.397)
log(Total assets)	-0.516	0.018	-0.001	0.328	0.171
	(0.494)	(0.137)	(0.034)	(0.305)	(0.257)
Return on assets, %	0.149^{***}	0.008	0.002	-0.057^{***}	-0.102^{***}
	(0.034)	(0.009)	(0.001)	(0.021)	(0.018)
Book to market ratio	-3.186^{***}	0.351^{*}	0.151^{**}	1.298^{***}	1.386^{***}
	(0.705)	(0.192)	(0.063)	(0.415)	(0.437)
Leverage	-0.553^{**}	-0.120^{*}	0.023	0.723^{***}	-0.072
	(0.236)	(0.067)	(0.015)	(0.154)	(0.096)
Year controls	Yes	Yes	Yes	Yes	Yes
Float and mk.cap. controls	Yes	Yes	Yes	Yes	Yes
Observations	29231	29231	29231	29231	29231
R^2	-0.380	0.434	-0.009	-0.416	-0.046
F stat.	805.7	1185.7	55.4	837.5	494.6

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

blindly follow ISS or firm's management recommendations. Moreover, this observation also suggests that firms likely have some market power since the Fisher separation theorem does not seem to hold.

I discover that other shareholders react to the mutual funds' portfolio structure. When a firm is held by mutual funds with closely overlapping portfolios, other shareholders reduce the number of votes they cast. Rational apathy of investors is a plausible explanation here. This result also demonstrates that horizontal shareholding has a tangible effect on corporate governance process.

Appendix C

Variables dictionary

Table 3.9: Detailed definitions of variables used.

Variable	Definition	Data source
"Abstain" votes, %	The share of votes "Abstain" out of total shares outstanding. Measured in %.	ISS Voting Analytics
"Against"/"Withhold" votes, %	The share of votes "Against"/"Withhold" (aggregated by ISS) out of total shares outstanding. Measured in %.	ISS Voting Analytics
"For" votes, %	The share of votes "For" out of total shares outstanding. Measured in %.	ISS Voting Analytics
ln(Total assets)	Natural logarithm of firm's total assets.	Compustat
1 index fund	A dummy variable equal to 1 if one (and only one) mutual fund in a pair is an index fund (has flag "D" in the `findex_fund_flag` field).	CRSP Mutual Funds Database
2 index funds	A dummy variable equal to 1 if both mutual funds in a pair are index funds (have flag "D" in the `findex_fund_flag` field).	CRSP Mutual Funds Database
Banded state $\times Russell2000_t$	The interaction term between banded state and $Russell2000_t$	Bloomberg
Banded state	A dummy variable equal to 1, if a firm was in the banded region during the current year's reconstitution process (if the election date is past the the index constituents announcement date; or previous year if the election date is before that date).	Bloomberg

Variable	Definition	Data source
Book to market ratio	Book to market ratio of the firm. A computation resulting in a negative book to market ratio is treated as a missing value.	Compustat
Broker "Non-Votes", %	The portion of shares non-voted (no vote cast) by a broker out of total shares outstanding. Measured in %.	ISS Voting Analytics
CIK	Central Index Key used by the SEC	Edgar
Expense ratio	An absolute or geometric average of mutual funds' expense ratios. Data is as of the most recently completed fiscal year. When geometric average is computed, both ratios are censored at zero if negative, multiplied, and then a square root is taken. For absolute difference no censoring is applied. Final result is converted to %.	CRSP Mutual Funds Database
Family size	A natural logarithm of a geometric average or an absolute difference of mutual funds' families sizes. Family size is computed by adding all total net assets of funds belonging to a family. ISS Voting Analytics provides family structure. CRSP Mutual Fund Database provides funds' total net assets.	ISS Voting Analytics and CRSP Mutual Fund Database
Firm Leverage	Firm Leverage computed from the Compustat data.	Compustat
Fund turnover ratio	An absolute or geometric average of mutual funds' turnover ratios. When geometric average is computed, both ratios are censored at zero if negative, multiplied, and then a square root is taken.	CRSP Mutual Funds Database
ISS Against another item	A dummy variable equal to 1 if ISS issues a recommendation to vote against another director nominee at the meeting.	ISS Voting Analytics
ISS Recommendation "For"	A dummy variable equal to 1 if ISS recommends to vote "For" on the proposal, and 0 otherwise.	ISS Voting Analytics

Variable	Definition	Data source
Leverage	Firm's leverage. Computation results are truncated within an interval [-0.01, 10].	Compustat
Management fee	An absolute or geometric average of mutual funds' management fees. The ratio of management fee and average net assets. When geometric average is computed, both values are censored at zero if negative, multiplied, and then a square root is taken. For absolute difference the individual values below -3 are censored at -3.	CRSP Mutual Funds Database
NPXFileID	Name of the N-PX Form file from the SEC that contains data on the mutual funds votes. Used to match the ISS fund ids with additional data available in the N-PX form.	ISS Voting Analytics
Other "Non-Votes", %	The portion of shares non-voted (no vote cast) not by a broker out of total shares outstanding. Measured in %.	ISS Voting Analytics
$\text{R}ussell2000_t$	An indicator that a company belongs to Russell 2000 index this year if the election date is past the index constituents announcement date; or previous year (if the election date is before that date).	Bloomberg
$\text{R}ussell2000_{t-1}$	An indicator that a company belonged to Russell 2000 index last year if the election date is past the current year's index constituents announcement date; or two years ago (if the election date is before that date).	Bloomberg
Ratio of expense ratios	Ratio of mutual funds' expense ratios. Evaluated as the larger value divided by the smaller value (as order in a pair of mutual funds should not matter). Data is as of the most recently completed fiscal year. If smaller expense ratio is negative, the value is treated as missing.	CRSP Mutual Funds Database
Return on assets	Return on assets (firm)	Compustat
S&P 500	A dummy variable equal to 1 if firm is a constituent of S&P 500 index at the date of the election.	ISS Voting Analytics and Compustat

Variable	Definition	Data source
Same MSA	A dummy variable equal to 1 if both funds in a pair have their management company addresses within the same Metropolitan Statistical Area.	CRSP Mutual Funds Database
Same family	A dummy variable equal to 1 if both mutual funds belong to the same family.	ISS Voting Analytics
Total net assets	A natural logarithm of a geometric average or an absolute difference of mutual funds' total net assets. The raw values of funds' total net assets are censored at $0.1 if they are smaller than this threshold.	CRSP Mutual Funds Database
% of Total Equity	An absolute or geometric average of mutual funds' investments in the firm as percentages of total firm's equity.	CRSP Mutual Funds Database and Compustat
% of Total net assets	An absolute or geometric average of mutual funds' percentages of their portfolios invested in the firm. When geometric average is computed, both values are censored at zero if negative, multiplied, and then a square root is taken.	CRSP Mutual Funds Database
% owned by index funds	The portion of shares owned by index funds (have flag "D" in the `findex_fund_flag` field) out of total shares outstanding. Measured by aggregating all shares that index funds own at the company.	CRSP Mutual Funds Database
% owned by non-index funds	The portion of shares owned by non-index funds (do not have flag "D" in the `findex_fund_flag` field) out of total shares outstanding. Measured by aggregating all shares that non-index funds own at the company.	CRSP Mutual Funds Database

Data matching procedure

The ISS Voting Analytics dataset lacks a mutual fund identification variable that would be common with other popular datasets on mutual funds, like CRSP Mutual Funds database and Thomson Reuters 13f. This is a known issue in the literature, and Schwartz-Ziv & Wermers (2019); Matvos & Ostrovsky (2008) and Iliev & Lowry (2015) provide their solutions to the problem. In this paper, I improve on the combined approach by semi-manually verifying the funds' names match between the ISS's and SEC EDGAR's data.

First, I use the `NPXFileID` field to retrieve the corresponding file from EDGAR database for each record in Voting Analytics database. This allows me to associate a `CIK` field from EDGAR with voting records. Then, I focus on a subset of mutual funds with a same `NPXFileID` value and establish a match by funds' names between Voting Analytics and EDGAR file (I use `Series_Name` field from the N-PX filing). This step appends Voting Analytics data with `Series_ID` and `Ticker` fields that identify individual fund in an N-PX filing.

I perform name matching between funds within an N-PX filing (identified by `Series_Name`) and funds in ISS Voting Analytics dataset with a corresponding link to the N-PX file. I do so in a two step procedure. First, for a fund from ISS dataset I rank all funds from an N-PX filing by their Levenshtein distance in their names to the fund in question. For best matches with Levenshtein distance of 3 or smaller (where 0 corresponds to a perfect match) I assume that I assume that funds in both datasets represent the same fund. Second, for all unmatched funds (with minimum distance of 4 and larger) I conduct a manual name match (assisted by sorting N-PX filing's funds by their similarity to a fund in question). If no match seems reasonable, I assign a no-match label.

Second, I use ticker data from N-PX filings to match individual funds to CRSP Mutual Funds database. Since a ticker might be shared by different funds over time, I only accept matches that happen no more than 1 calendar year apart.

An alternative approach would be to use `crsp_cik_map` provided by WRDS. This linking dataset contains association between `Series_ID` and `CIK` fields from N-PX filings and corresponding `crsp_fundno`.

Finally, I use MFLinks dataset to connect CRSP Mutual Fund data to information in Thomson Reuters Mutual Fund Ownership dataset.

Bibliography

Abadie, A., A. Diamond, & J. Hainmueller (2010) Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association* **105**, 493–505.

Admati, A. R. & P. Pfleiderer (2009) The "Wall Street Walk" and Shareholder Activism: Exit as a Form of Voice. *Review of Financial Studies* **22**, 2645–2685.

Aggarwal, R., S. Dahiya, & N. R. Prabhala (2019) The power of shareholder votes: Evidence from uncontested director elections. *Journal of Financial Economics* **133**, 134–153.

Alesina, A. & H. Rosenthal (1996) A Theory of Divided Government. *Econometrica* **64**, 1311–1341.

Aliprantis, C. D., K. C. Border, & SpringerLink (Online service) (2006) *Infinite Dimensional Analysis [Electronic Resource] : A Hitchhiker's Guide / by Charalambos D. Aliprantis, Kim C. Border.* Springer-Verlag Berlin Heidelberg.

Allen, C. H. (2007) Study of Majority Voting in Director Elections. *SSRN Electronic Journal*.

Amjad, M. J., D. Shah, & D. Shen (2018) Robust Synthetic Control. *Journal of Machine Learning Research* **19**, 1–51.

Appel, I. R., T. A. Gormley, & D. B. Keim (2016) Passive investors, not passive owners. *Journal of Financial Economics* **121**, 111–141.

Appel, I. R., T. A. Gormley, & D. B. Keim (2018) Standing on the shoulders of giants: The effect of passive investors on activism. *The Review of Financial Studies* **32**, 2720–2774.

Arrow, 1921-2017, K. J. K. J. (1963) *Social Choice and Individual Values.* Wiley.

Atanasov, V. & B. Black (2016) Shock-Based Causal Inference in Corporate Finance and Accounting Research. *Critical Finance Review* **5**, 207–304.

Avis, D., P. Bose, G. Toussaint, T. Shermer, B. Zhu, & J. Snoeyink (1996) On the sectional area of convex polytopes. *Proceedings of the twelfth annual symposium on computational geometry*, 411–412.

Azar, J. (2017) Portfolio Diversification, Market Power, and the Theory of the Firm. *SSRN Electronic Journal*.

Backus, M., C. Conlon, & M. Sinkinson (2019) The Common Ownership Hypothesis: Theory and Evidence. *Brookings working paper*.

Backus, M., C. Conlon, & M. Sinkinson (2019) Common Ownership in America: 1980-2017. *NBER Working Paper No. 25454*, 56.

Baig, A., J. DeLisle, & G. R. Zaynutdinova (2018) Passive Ownership and Earnings Manipulation. *SSRN Electronic Journal*.

Banks, J. S. & J. Duggan (2005) Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In *Social Choice and Strategic Decisions*, pp. 15–56. Springer.

Banzhaf, John F. III (1964) Weighted Voting Doesn't Work: A Mathematical Analysis. *Rutgers Law Review* **19**, 317–343.

Barua, A., K. Raghunandan, & D. V. Rama (2017) Shareholder Votes on Auditor Ratification and Subsequent Auditor Dismissals. *Accounting Horizons* **31**, 129–139.

Bebchuk, L. A. & S. Hirst (2019) The Specter of the Giant Three. *SSRN Electronic Journal*.

Benninga, S. & E. Muller (1979) Majority Choice and the Objective Function of the Firm under Uncertainty. *The Bell Journal of Economics* **10**, 670–682.

Berge, C. (1963) *Topological Spaces, Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity. Tr. by E. M. Patterson.* Oliver & Boyd.

Berger, E. A., A. W. Butler, E. Hu, & M. Zekhnini (2020) Financial integration and credit democratization: Linking banking deregulation to economic growth. *Journal of Financial Intermediation*, 100857.

Berger, M. (1987) *Geometry* Number v. 2 in Universitext (1979). Springer-Verlag.

Bohlin, L. & M. Rosvall (2014) Stock Portfolio Structure of Individual Investors Infers Future Trading Behavior. *PLoS ONE* **9**, e103006.

Bolton, P., T. Li, E. Ravina, & H. Rosenthal (2018) Investor Ideology. *SSRN Electronic Journal*.

Boone, A. L. & J. T. White (2015) The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics* **117**, 508–533.

Brav, A., W. Jiang, T. Li, & J. Pinnington (2019) Picking Friends Before Picking (Proxy) Fights: How Mutual Fund Voting Shapes Proxy Contests. *SSRN Electronic Journal*.

Brito, D., E. Elhauge, R. Ribeiro, & H. Vasconcelos (2019) Modeling Horizontal Shareholding with Ownership Dispersion. *Economics Working Papers, Católica Porto Business School, Universidade Católica Portuguesa* **01**, 17.

Brito, D., A. Osório, R. Ribeiro, & H. Vasconcelos (2018) Unilateral effects screens for partial horizontal acquisitions: The generalized HHI and GUPPI. *International Journal of Industrial Organization* **59**, 127–189.

Bubb, R. & E. Catan (2018) The Party Structure of Mutual Funds. *SSRN Electronic Journal*.

Cai, J., J. L. Garner, & R. A. Walkling (2009) Electing Directors. *The Journal of Finance* **64**, 2389–2421.

Cai, J., J. L. Garner, & R. A. Walkling (2013) A paper tiger?An empirical analysis of majority voting. *Journal of Corporate Finance* **21**, 119–135.

Cha, S.-H. (2007) Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions. *International Journal Of Mathematical Models And Methods In Applied Sciences* **1**, 8.

Chang, Y.-C., H. Hong, & I. Liskovich (2015) Regression Discontinuity and the Price Effects of Stock Market Indexing. *Review of Financial Studies* **28**, 212–246.

Choi, S., J. E. Fisch, & M. Kahan (2010) The Power of Proxy Advisors: Myth or Reality?*Emory Law Journal* **59**, 869–918.

Choi, S. J., J. E. Fisch, & M. Kahan (2009) Director elections and the role of proxy advisors. *Southern California Law Review* **82**, 649–702.

Coates, J. C. (2018) The Future of Corporate Governance Part I: The Problem of Twelve. *SSRN Electronic Journal*.

Cotter, J. F., A. R. Palmiter, & R. S. Thomas (2013) The First Year of 'Say on Pay' Under Dodd-Frank: An Empirical Analysis and Look Forward. *Geo. Wash. L. Rev* **81**, 2013.

Council of Institutional Investors (2017) FAQ: Majority Voting for Directors. *Website Brochure*, 1–10.

Crès, H. & M. Tvede (2005) Portfolio diversification and internalization of production externalities through majority voting. *<hal-01065579>*.

Cuñat, V., M. Giné, & M. Guadalupe (2016) Say Pays! Shareholder Voice and Firm Performance. *Review of Finance* **20**, 1799–1834.

Cvijanovic, D., M. Groen-Xu, & K. E. Zachariadis (2019) Free-Riders and Underdogs: Participation in Corporate Voting. *SSRN Electronic Journal*.

Dao, M., S. Mishra, & K. Raghunandan (2008) Auditor Tenure and Shareholder Ratification of the Auditor. *Accounting Horizons* **22**, 297–314.

Davis, G. F. (2008) A new finance capitalism?Mutual funds and ownership re-concentration in the United States. *European Management Review* **5**, 11–21.

DeAngelo, H. (1981) Competition and Unanimity. *The American Economic Review* **71**, 18–27.

DeMarzo, P. M. (1993) Majority Voting and Corporate Control: The Rule of the Dominant Shareholder. *The Review of Economic Studies* **60**, 713.

Elhauge, E. R. (2019) The Causal Mechanisms of Horizontal Shareholding. *SSRN Electronic Journal*.

Elhauge, E. R. (2019). The Greatest Anticompetitive Threat of Our Time: Fixing the Horizontal Shareholding Problem. *ProMarket*.

Elhauge, E. R. (2019) How Horizontal Shareholding Harms Our Economy - And Why Antitrust Law Can Fix It. *SSRN Electronic Journal*.

Enriques, L. & A. Romano (2018) Institutional Investor Voting Behavior: A Network Theory Perspective. *SSRN Electronic Journal*.

Ertimur, Y., F. Ferri, & D. Oesch (2013) Shareholder Votes and Proxy Advisors: Evidence from Say on Pay. *Journal of Accounting Research* **51**, 951–996.

European Commission (2017). CASE M.7932 – Dow/DuPont.

Fama, E. F. & A. B. Laffer (1972) The Number of Firms and Competition. *The American Economic Review* **62**, 670–674.

Feddersen, T. J. & W. Pesendorfer (1996) The Swing Voter's Curse. *The American Economic Review* **86**, 408–424.

Fichtner, J., E. M. Heemskerk, & J. Garcia-Bernardo (2017) Hidden power of the Big Three?Passive index funds, re-concentration of corporate ownership, and new financial risk. *Business and Politics* **19**, 298–326.

Fischer, P. E., J. D. Gramlich, B. P. Miller, & H. D. White (2009) Investor perceptions of board performance: Evidence from uncontested director elections. *Journal of Accounting and Economics* **48**, 172–189.

Fisher, I. (1930) *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It* New York: The Macmillan Company.

Fos, V. & M. Tsoutsoura (2014) Shareholder Democracy in Play: Career Consequences of Proxy Contests. *Journal of Financial Economics* **114**, 316–340.

FTSE Russell (2019). Construction and Methodology of Russell U.S. Equity Indexes.

Getmansky, M., G. Girardi, K. Hanley, S. Nikolova, & L. Pelizzon (2018) Portfolio Similarity and Asset Liquidation in the Insurance Industry. *SSRN Electronic Journal*, 25.

Gloßner, S. (2018) The Effects of Institutional Investors on Firm Outcomes: Empirical Pitfalls of Quasi-Experiments Using Russell 1000/2000 Index Reconstitutions. *SSRN Electronic Journal*.

GMI Ratings (2013). Vote Calculation Methodologies. Technical report, GMI Ratings.

Gordon, R. H. (2003) Do Publicly Traded Corporations Act in the Public Interest? *The B.E. Journal of Economic Analysis & Policy* **3**.

Grossman, S. J. & J. E. Stiglitz (1977) On Value Maximization and Alternative Objectives of the Firm. *The Journal of Finance* **32**, 389.

Grundfest, J. A. (1993) Just Vote No: A Minimalist Strategy for Dealing with Barbarians inside the Gates. *Stanford Law Review* **45**, 857.

Guercio, D. D., L. Seery, & T. Woidtke (2008) Do boards pay attention when institutional investor activists "just vote no"? *Journal of Financial Economics* **90**, 84–103.

Hanley, K. W. & G. Hoberg (2010) The Information Content of IPO Prospectuses. *Review of Financial Studies* **23**, 2821–2864.

Hanley, K. W. & G. Hoberg (2012) Litigation risk, strategic disclosure and the underpricing of initial public offerings. *Journal of Financial Economics* **103**, 235–254.

Hansen, R. G. & J. R. Lott (1996) Externalities and Corporate Objectives in a World with Diversified Shareholder/Consumers. *The Journal of Financial and Quantitative Analysis* **31**, 43–68.

Harford, J., D. Jenter, & K. Li (2011) Institutional cross-holdings and their effect on acquisition decisions. *Journal of Financial Economics* **99**, 27–39.

Harris, M. & A. Raviv (1988) Corporate governance. Voting rights and majority rules. *Journal of Financial Economics* **20**, 203–235.

Hart, O. D. (1979) On Shareholder Unanimity in Large Stock Market Economies. *Econometrica* **47**, 1057–1083.

Hayashi, F. (2000) *Econometrics* Princeton University Press.

He, J., J. Huang, & S. Zhao (2019) Internalizing Governance Externalities: The Role of Institutional Cross-Ownership. *Journal of Financial Economics*.

Heath, D., D. Macciocchi, R. Michaely, & M. C. Ringgenberg (2018) Passive Investors are Passive Monitors. *Working paper*, 49.

Iliev, P., K. V. Lins, D. P. Miller, & L. Roth (2015) Shareholder Voting and Corporate Governance Around the World. *The Review of Financial Studies* **28**, 2167–2202.

Iliev, P. & M. Lowry (2015) Are Mutual Funds Active Voters? *Review of Financial Studies* **28**, 446–485.

Iliev, P. & S. Vitanova (2019) The Effect of the Say-on-Pay Vote in the United States. *Management Science* **65**, 4505–4521.

Jayaraman, N., A. Khorana, & E. Nelling (2002) An Analysis of the Determinants and Shareholder Wealth Effects of Mutual Fund Mergers. *The Journal of Finance* **57**, 1521–1551.

Jill E. Fisch (2017) Standing Voting Instructions: Empowering the Excluded Retail Investor. *Minnesota Law Review* **82**, 11–60.

Khachiyan, L. (1993) Complexity of Polytope Volume Computation. In *New Trends in Discrete and Computational Geometry*. Springer.

Khorana, A., P. Tufano, & L. Wedge (2007) Board Structure, Mergers and Shareholder Wealth: A Study of the Mutual Fund Industry. *Journal of Financial Economics* **85**, 571–598.

Krishnan, J. & Z. S. Ye (2005) Why Some Companies Seek Shareholder Ratification on Auditor Selection. *Accounting Horizons* **19**, 237–254.

Kwon, O.-W. & J.-H. Lee (2003) Text categorization based on k-nearest neighbor approach for Web site classification. *Information Processing & Management* **39**, 25–44.

Lafarre, A. (2017) Small Shareholder Participation in European AGMs. *SSRN Electronic Journal*.

Lasserre, J. B. (1983) An Analytical Expression and an Algorithm for the Volume of a Convex Polyhedron in R“. *Journal of Optimization Theory and Applications* **39**, 15.

Lasserre, J. B. (1998) Integration On a Convex Polytope. *Proceedings of the American Mathematical Society* **126**, 2433–2441.

Li, Y. & D. Yermack (2016) Evasive shareholder meetings. *Journal of Corporate Finance* **38**, 318–334.

Lindsey, L. (2008) Blurring Firm Boundaries: The Role of Venture Capital in Strategic Alliances. *The Journal of Finance* **63**, 1137–1168.

Malenko, A. & N. Malenko (2019) Proxy Advisory Firms: The Economics of Selling Information to Voters. *Journal of Finance* **74**, 2441–2490.

Malenko, N. & Y. Shen (2016) The Role of Proxy Advisory Firms: Evidence from a Regression-Discontinuity Design. *Review of Financial Studies* **29**, 45.

Matt Egan (2014). Just 27% of investors bother to vote - The Buzz - Investment and Stock Market News. *CNN Website*. http://buzz.money.cnn.com/2014/06/12/shareholders-dont-vote/.

Matvos, G. & M. Ostrovsky (2008) Cross-ownership, returns, and voting in mergers. *Journal of Financial Economics* **89**, 391–403.

McLemore, P. (2019) Do Mutual Funds Have Decreasing Returns to Scale? Evidence from Fund Mergers. *Journal of Financial and Quantitative Analysis* **54**, 1683–1711.

Milne, F. (1974) Corporate investment and finance theory in competitive equilibrium. *Economic Record* **50**, 511–533.

Milne, F. (1981) The firm's objective function as a collective choice problem. *Public Choice* **37**, 473–486.

Monks, R. A. G. & N. Minow (2004) *Corporate Governance* (3rd ed.). Blackwell Pub.

Morton, F. M. S. & H. J. Hovenkamp (2018) Horizontal Shareholding and Antitrust Policy. *The Yale Law Journal* **127**, 2026–2047.

Mullins, W. (2014) The Governance Impact of Index Funds: Evidence from Regression Discontinuity. *Working paper*, 60.

Myronenko, A. & X. Song (2010) Point-Set Registration: Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**, 2262–2275.

Nili, Y. & K. Kastiel (2016) In Search of "Absent" Shareholders: A New Solution To Retail Investors' Apathy. *Delaware Journal of Corporate Law* **41**, 55–104.

Robin Ferracone & Dayna Harris (2011) Say On Pay: Identifying Investor Concerns. *Council of Institutional Investors*, 1–32.

Rotemberg, J. (1984) Financial transaction costs and industrial performance. *Sloan School of Management*, 1554–84.

Sainty, B. J., G. K. Taylor, & D. D. Williams (2002) Investor Dissatisfaction toward Auditors. *Journal of Accounting, Auditing & Finance* **17**, 111–136.

Salop, S. & D. O'Brien (2000) Competitive Effects of Partial Ownership: Financial Interest and Corporate Control. *Antitrust L.J.* **67**, 559–614.

Sanderson, E. & F. Windmeijer (2016) A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics* **190**, 212–221.

Schmalz, M. C. (2018) Common-Ownership Concentration and Corporate Conduct. *Annual Review of Financial Economics* **10**, 413–448.

Schmidt, C. & R. Fahlenbrach (2017) Do exogenous changes in passive institutional ownership affect corporate governance and firm value? *Journal of Financial Economics* **124**, 285–306.

Schnell, D. K. & A. Chen (2019) Counting Shareholder Votes. *Insights: The Corporate & Securities Law Advisor* **33**, 3–7.

Schwartz-Ziv, M. & R. Wermers (2019) Do Institutional Investors Monitor Their Large vs. Small Investments Differently? Evidence from the Say-on-Pay Vote. *SSRN Electronic Journal*.

Sebastiani, F. (2002) Machine Learning in Automated Text Categorization. *ACM Computing Surveys* **34**, 1–47.

SEC (2003). Final Rule: Disclosure of Proxy Voting Policies and Proxy Voting Records by Registered Management Investment Companies, 17 CFR Parts 239, 249, 270, and 274.

SEC (2003) Standards Relating To Listed Company Audit Committees. *Securities and Exchange Commission.*

SEC (2009) Order Approving Proposed Rule Change. *Securities and Exchange Commission.*

Sias, R., H. J. Turtle, & B. Zykaj (2013) Hedge Fund Crowds and Mispricing. *Management Science* **62**, 764–784.

Stock, J.H. (2005) Testing for weak instruments in linear IV regression. In *Identification and Inference for Econometric Models, Essays in Honor of Thomas Rothenberg.*, pp. 557–586. Cambridge University Press.

Stokdyk, P. S. B. & J. H. Trotter (2016) Proxy Disclosure Recommendations. *Harvard Law School Forum on Corporate Governance and Financial Regulation*, 4.

Sundaram, R. K. (1996) *A First Course in Optimization Theory / Rangarajan K. Sundaram.* Cambridge University Press.

Tao, T. (2010) *An Epsilon of Room: Pages from Year Three of a Mathematical Blog* AMS Press, Inc.

Tao, T. (2011) *An Introduction to Measure Theory*Volume 126 of *Graduate Studies in Mathematics.* American Mathematical Society.

The Office of Investor Education and Advocacy (2012). Spotlight on Proxy Matters - The Mechanics of Voting. http://www.sec.gov/spotlight/proxymatters/voting_mechanics.shtml.

Van der Elst, C. (2011) Revisiting Shareholder Activism at AGMs: Voting Determinants of Large and Small Shareholders. *SSRN Electronic Journal.*

Wei, W. & A. Young (2017) Selection Bias or Treatment Effect? A Re-Examination of Russell 1000/2000 Index Reconstitution. *SSRN Electronic Journal.*

Yermack, D. (2010) Shareholder Voting and Corporate Governance. *Annual Review of Financial Economics* **2**, 103–125.

Zhao, X. (2005) Exit Decisions in the U.S. Mutual Fund Industry. *The Journal of Business* **78**, 1365–1402.