

Free-riders and Underdogs: Participation in Corporate Voting

Dragana Cvijanović
UNC Chapel Hill

Moqi Groen-Xu
LSE

Konstantinos E. Zachariadis
Queen Mary University of London

ABFER 7th Annual Conference
May 2019

Issue

- ▶ Voting is the main avenue in which shareholder preferences are aggregated.
- ▶ Many investors do not have to vote in shareholder meetings (e.g., US: mutual funds vs. hedge funds).
- ▶ Yet many do. Discretionary participation from our data (i.e., US, S&P 1500 firms, 2003-2011):
 - ▶ on average 73%;
 - ▶ considerable variation between sponsors, proposal types, and firms.
- ▶ Does voting participation affect outcomes? And if yes by how much?

What We Do

- ▶ Develop a rational choice model where participation depends on the cost and benefit of voting and the probability that one's vote matters (i.e., a pivotal voter model).
- ▶ Innovation relative to political voting literature is ownership heterogeneity (regular vs discretionary voters).
- ▶ Then use the structure of the model to estimate unobserved shareholder preferences from aggregate US voting data.
- ▶ Conduct counterfactual analysis relative to the full participation benchmark and relative to scenarios with different costs of voting.

Model - Setup (I/II)

- ▶ Firm with $n + 1$ voting shares split between two *groups* of voters:
 - ▶ γ (fraction) *regular voters*, who always vote;
 - ▶ $1 - \gamma$ *discretionary voters*, who choose whether to vote;
 - ▶ Discretionary voters have a single share and hence a single vote each.
- ▶ Proposal is a vote between options R or L :
 - ▶ Shareholders are born with preference *types*, R or L ;
 - ▶ So they are 'partisan' or in a corporate context: 'disagree'.
- ▶ Voter preference types:
 - ▶ q of regular voters support R , while $(1 - q)$ support L ;
 - ▶ q is fully observed and wlog $q \in (1/2, 1)$;
 - ▶ In discretionary voters R has ex ante popularity $p \in (0, 1)$;
 - ▶ Crux of the model is that p is unknown;
 - ▶ So model features aggregate uncertainty.

Model - Setup (II/II)

- ▶ Discretionary voters decide to participate in voting based on:
 - ▶ benefit v , which they receive only if their preferred option wins vs
 - ▶ cost c , which they face when they vote, regardless of the outcome.
- ▶ Voters simultaneously and confidentially cast their votes; the outcome is decided by simple majority over the votes cast; in case of a tie, a fair coin toss is the tie-breaker.
- ▶ All the above are common knowledge. The only choice variable (strategy) is whether a discretionary voter votes.
- ▶ We look for symmetric (pure or mixed) strategies within types (R or L) and the solution concept is Bayesian Nash Equilibrium.

Model - Analysis

- ▶ A focal discretionary voter participates if $v \Pr[\text{Pivotal}|i] > c$, or

$$\Pr[\text{Pivotal}|i] > c/v, \text{ for } i \in \{R, L\}.$$

If it is smaller she does not participate, while if it is equal she is indifferent.

- ▶ To solve for the equilibrium we rely on calculations of the pivotal probabilities for large electorates (i.e., n large) similar to Myatt (2015).
- ▶ Let $t_i \in [0, 1]$ the discretionary participation rate for voter of type $i \in \{R, L\}$.
- ▶ Ruling out trivial equilibria (with $t_L = 0$ where R wins) there are 6 possible equilibria to compute $t_L \in \{(0, 1), 1\}$, $t_R \in \{0, (0, 1), 1\}$.

Model - Equilibria

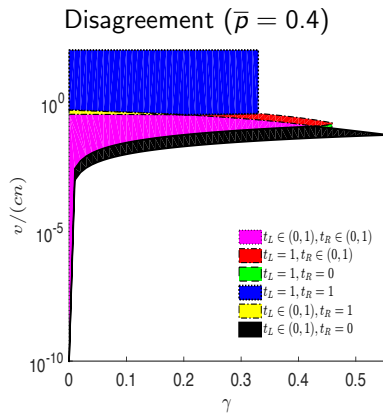
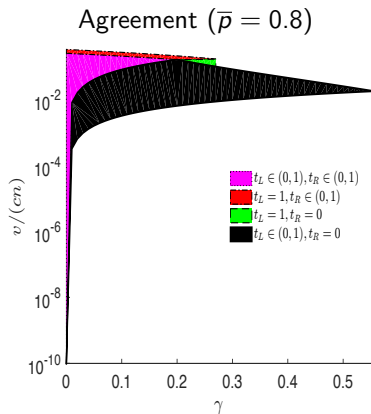
- ▶ We derive the parameter regions of the model $\{\gamma, v/(cn), q, l, h\}$ where each equilibrium exists, for $p \sim \mathcal{U}[l, h] \subseteq [0, 1]$.
- ▶ We also derive the possible outcome under each equilibrium:

Eqm	Participation		Avg outcome		
	t_L	t_R	L wins	Tie	R wins
<i>mm</i>	(0, 1)	(0, 1)		✓	
<i>1m</i>	1	(0, 1)		✓	✓
10	1	0			✓
11	1	1	✓	✓	✓
<i>m1</i>	(0, 1)	1	✓		
<i>m0</i>	(0, 1)	0			✓

- ▶ In all equilibria:
 - ▶ $\partial t_L / \partial q \geq 0$. This is the (intergroup) underdog effect.
 - ▶ $\partial t_R / \partial q \leq 0$. This is the (intergroup) free-riding effect.

Model - Equilibria Regions

We plot the *non-overlapping* equilibrium regions in the plane:
size of the regular block $\gamma \times$ benefit-cost ratio per voter $v/(cn)$.



$$h - l = 0.2, q = 0.9$$

Data

For 18,520 proposals in all S&P 1500 firms between 2003–2011:

- ▶ Aggregate voting data from ISS (excl. director election and routine proposals);
- ▶ Ownership data from 13F form filings and proxy statements;
- ▶ Direction of vote for Form N-PX filers.

Mandatory voting:

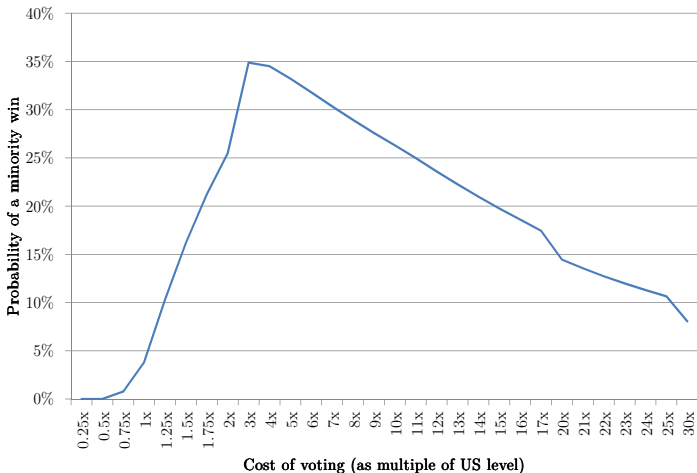
- ▶ Mutual funds and investment advisors have a fiduciary duty to vote and report their vote (in the N-PX forms);
- ▶ Very few (2%) non-institution blockholders over 5%;
- ▶ We use “N-PX filers” as an empirical approximation of “regular voters”.

Selection Effects in the US Data

- ▶ We use the observed in the data:
 - ▶ regular voter (i.e., N-PX) characteristics (γ, q) and discretionary support of both types amongst those who vote, let $dSuR, dSuL$, to estimate using GMM (with four moment conditions) [More on Estimation](#)
 - ▶ the benefit to cost ratio per voter $v / (cn)$, the avg and stdev of the fraction of discretionary voters supporting $R \bar{p}$ and $\text{std}(\rho)$.
- ▶ We find that most proposals correspond to the equilibrium with full participation by the underdog and partial participation by the other side (i.e., equilibrium $1m$).
- ▶ We use the estimated parameters to compare the observed voting outcomes from the benchmark of the counterfactual under full participation:
 - ▶ The more popular choice on avg (i.e., R) receives 21% less support;
 - ▶ The observed voting decision differs by an average 3.7%;
 - ▶ The probability of underdog wins is highest for governance related shr proposals and in general larger for shr than mgmt proposals.

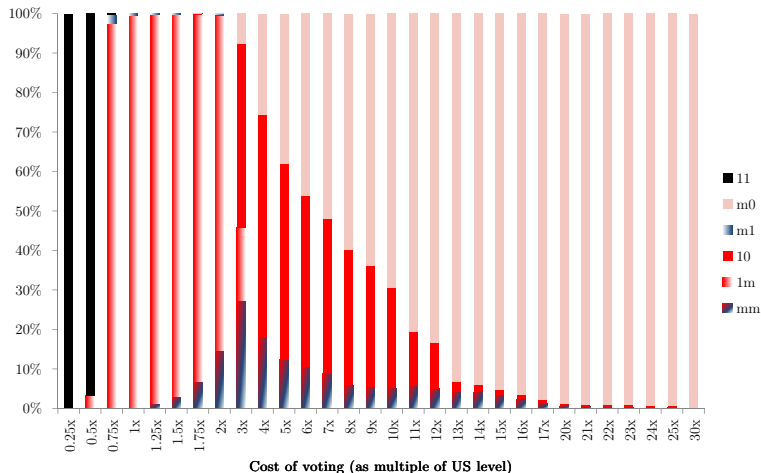
(More) Counterfactuals: Misalignment

- ▶ *Application:* recent regulatory attempts to cut EU cost of voting to US level (i.e., the Shareholder Rights Directive).
- ▶ Effects are non-linear: misalignment has a reverse U-shape in cost.



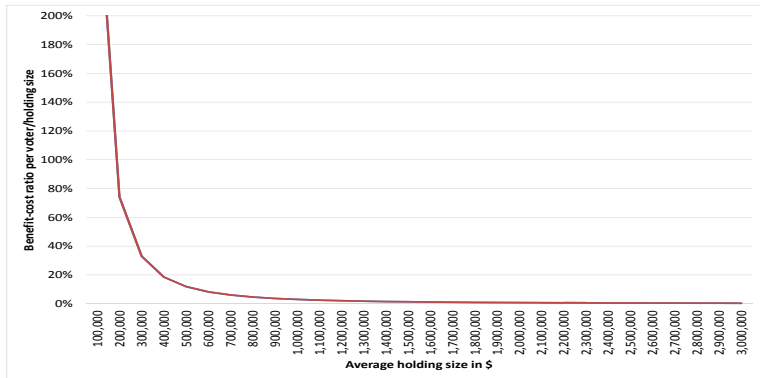
(More) Counterfactuals: Distribution of Equilibria

- ▶ US level: mostly equilibrium $1m$ (i.e., turning out to lose on avg).
- ▶ Peak probability at $3\times$ US level: the level with the most mm equilibrium (i.e., where the avg outcome is decided by a coin flip).



Benefit to Cost Ratio

- ▶ We plot the “return” of a proposal (assuming cost \$1) by dividing the benefit to cost ratio estimate by the assumed block size.
- ▶ For an avg share holding of \$1.5 million (see (Ahern (2015) for insiders) the “return” is 1.3% (Cuñat et'al (2012) find 1.6%).
- ▶ The benefit to cost ratio is highest for mgmt proposals on takeover defense; smallest for board and governance related shr proposals.



Estimation Performance

- ▶ The algorithm assigns 95% (under the single-stage baseline specification, 86% under two-stage GMM) of the sample proposals to an equilibrium.
- ▶ Our estimation significantly outperforms models from the previous literature (e.g., Malenko & Shen (2016)).
- ▶ Robust to alternative target moments, use of quantiles.
- ▶ Holds in various subsamples (ownership, equity lending supply/demand, number of proposals per meeting, information content).

Literature

- ▶ Voting participation in political elections: Palfrey & Rosenthal (1983), Feddersen & Pesendorfer (1996,1997), Feddersen (2004), Krishna & Morgan (2011,2012), Evren (2012), Myatt (2015).
- ▶ Corporate voting theory: Maug & Rydqvist (2008), Levit & Malenko (2011), Bar-Isaac & Shapiro (2017).
- ▶ Empirical work on the importance of shareholder proposals: Cuñat et al. (2012), Metzger & Bach (2017).
- ▶ Empirical work on mutual fund voting (N-PX): Brickley et al. (1994), Matvos & Ostrovsky (2006), Cvijanović et al. (2015), Malenko & Shen (2016).
- ▶ Empirical work on corporate voting participation: Van der Elst (2011).
- ▶ Recent (theoretical and empirical) work on shareholder disagreement: Bolton et'al (2018), Li et'al (2019).
- ▶ Recent empirical work on retail shareholder voting: Brav et'al (2019).

Conclusions

- ▶ Model that links observed participation rates with unobserved characteristics of shareholder preferences:
 - ▶ *Free-riding effect*: Agreement yields lower participation rates;
 - ▶ *Underdog effect*: Disagreement yields higher participation rates.
- ▶ Using the model and US data we structurally estimate underlying shareholder preferences.
- ▶ Document large selection effects towards the “underdog”.
- ▶ Equilibrium regions that inform counterfactual analysis: three times the U.S. level cost of voting corresponds to 35% of misalignment relative to the full participation case.

APPENDIX

Estimation Algorithm I/II

- ▶ First, create the bins: we sort our data into quantiles of γ , quantiles of n , and proposal-types. For each bin we compute: the averages $\bar{\gamma}$ and \bar{q} ; the averages \overline{dSuL} , \overline{dSuR} , $\overline{dSuL^2}$, and $\overline{dSuR^2}$;
- ▶ Second, exhaustive search in the space $\{v/(cn), \bar{p}, \text{std}(p)\}$ and for each point in the grid and each possible equilibrium:
 - i) calculate the interval Γ and ask if $\bar{\gamma} \in \Gamma$, if yes continue, o/w proceed to the following equilibrium;
 - ii) if $\bar{\gamma} \in \Gamma$ then calculate the interval V and ask if $v/(cn) \in V$, if yes continue, o/w proceed to the following equilibrium;
 - iii) if $v/(cn) \in V$ then calculate t_L and t_R and create estimates for \overline{dSuL}_{est} , \overline{dSuR}_{est} , $\overline{dSuL^2}_{est}$, $\overline{dSuR^2}_{est}$;
 - iv) finally, calculate the estimation error:

$$\begin{aligned} \text{Estimation Error} &= (\overline{dSuL}_{est} - \overline{dSuL})^2 + (\overline{dSuR}_{est} - \overline{dSuR})^2 \\ &+ (\overline{dSuL^2}_{est} - \overline{dSuL^2})^2 + (\overline{dSuR^2}_{est} - \overline{dSuR^2})^2; \end{aligned}$$

- ▶ Third, pick the point in the grid and associated equilibrium that minimizes the error.

Estimation Algorithm II/II

- ▶ Our *identifying assumption* is that within each bin (i.e., a quintile of γ , quintile of n , and proposal-type) unobserved $\{v/(cn), \bar{p}, \text{std}(p)\}$ are constant and the averages $\bar{\gamma}, \bar{q}$ are representative. Hence, variation in (discretionary support for R) p across proposals is the (only) variation that allows us to identify the bin-specific parameters.
- ▶ The algorithm returns a point estimate for $v/(nc)$ only if in the estimated equilibrium not both rates are “corner” (i.e., equilibria $mm, m1, m0, 1m$). Otherwise (i.e., equilibria 11 and 10), we obtain a set estimate: $\{v/(nc)_{\text{lower}}, v/(nc)_{\text{upper}}\}$.
- ▶ Given uniqueness of equilibrium for specific parameter values we can be certain that no other (estimated) parameter values (within the grid) and corresponding equilibrium result in lower estimation error given the observed data.

Delta Method

We compute our standard errors using the Delta Method approach:

- ▶ First, we find numerically $\partial\theta_i/\partial m_j$, where $\theta = [v/(nc)_{\text{lower}}, v/(nc)_{\text{upper}}, \bar{p}, \text{std}(p)]$, $m = [\overline{dSuL}, \overline{dSuR}, \overline{dSuL^2}, \overline{dSuR^2}]$, so $i, j \in \{1, 2, 3, 4\}$.
- ▶ Second, we estimate the variance-covariance matrix, let S , of the four errors that we base our estimation on

$$\begin{array}{cc} \underbrace{\overline{dSuL}_{\text{est}}}_{t_L(1-\bar{p})} - dSuL, & \underbrace{\overline{dSuR}_{\text{est}}}_{t_R\bar{p}} - dSuR, \\ \underbrace{\overline{dSuL^2}_{\text{est}}}_{(t_L\text{std}(p))^2 + (t_L(1-\bar{p}))^2} - dSuL^2, & \underbrace{\overline{dSuR^2}_{\text{est}}}_{(t_R\text{std}(p))^2 + (t_R\bar{p})^2} - dSuR^2, \end{array}$$

where $t_L, t_R, \bar{p}, \text{std}(p)$ are estimates.

- ▶ Third, the variance of our error in estimating parameter θ_i is

$$\Delta_i \times S \times \Delta_i^T,$$

where vector $\Delta_i \equiv [\partial\theta_i/\partial m_1, \partial\theta_i/\partial m_2, \partial\theta_i/\partial m_3, \partial\theta_i/\partial m_4]$, for $i = \{1, 2, 3, 4\}$.

Misalignment Method

We compute the probability of misalignment *per bin* as follows:

- ▶ Given our estimated \bar{p} , $\text{std}(p)$ we simulate proposals $p \sim \mathcal{U}[l, h]$.
- ▶ For each p given our estimated t_L, t_R for this bin we compute:
 - ▶ The estimated outcome index under discretionary participation:

$$O_{disc}(p) \equiv \left[\underbrace{\bar{\gamma}\bar{q} + (1 - \bar{\gamma}) t_R p}_{\text{support for } R} \right] - \left[\underbrace{\bar{\gamma}(1 - \bar{q}) + (1 - \bar{\gamma}) t_L(1 - p)}_{\text{support for } L} \right].$$

- ▶ The estimated outcome index under full participation is:

$$O_{full}(p) \equiv \left[\underbrace{\bar{\gamma}\bar{q} + (1 - \bar{\gamma}) p}_{\text{support for } R} \right] - \left[\underbrace{\bar{\gamma}(1 - \bar{q}) + (1 - \bar{\gamma})(1 - p)}_{\text{support for } L} \right].$$

- ▶ And whether they differ in their assigned outcome

$$\mathbb{I}(O_{disc}(p)O_{full}(p) \leq 0).$$

- ▶ Finally, we average for all p and this gives as a per bin estimate of $\mathbb{P}[O_{disc}(p)O_{full}(p) \leq 0]$.