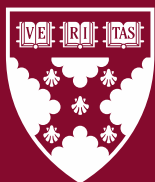


Working Paper 21-131

Catching Outliers: Committee Voting and the Limits of Consensus when Financing Innovation

Andrey Malenko
Ramana Nanda
Matthew Rhodes-Kropf
Savitar Sundaresan



**Harvard
Business
School**

Catching Outliers: Committee Voting and the Limits of Consensus when Financing Innovation

Andrey Malenko

University of Michigan

Ramana Nanda

Imperial College London

Harvard Business School

Matthew Rhodes-Kropf

Massachusetts Institute of Technology

Savitar Sundaresan

Imperial College London

Working Paper 21-131

Copyright © 2021, 2022, 2023 Andrey Malenko, Ramana Nanda, Matthew Rhodes-Kropf, and Savitar Sundaresan.

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Funding for this research was provided in part by Harvard Business School. Nanda is Professor of Entrepreneurial Finance and Academic Lead of the Institute for Deep Tech Entrepreneurship at Imperial College London and Visiting Scholar at Harvard Business School. Rhodes-Kropf is an advisor to, and investor in, Correlation Ventures, which provided some of the data. This paper is part of a project that has received funding from the European Research Council (ERC) Horizon 2020 research and innovation programme, Grant Number 865127.

Catching Outliers: Committee Voting and the Limits of Consensus when Financing Innovation*

ANDREY MALENKO

Boston College

RAMANA NANDA

Imperial College London

MATT RHODES-KROPF

MIT Sloan

SAVITAR SUNDARESAN

Imperial College London

November 2023

Abstract

We document that investment committees of major VCs use a voting rule where one partner ‘championing’ an early-stage investment is sufficient to invest. Their stated reason for this rule is to ‘catch outliers’. The same VCs use a more conventional ‘majority’ rule for later-stage investments. This evidence points to a model in which voting partners get signals about different project dimensions and superstar projects excel on some dimensions even if flawed on others. In this case, if the distribution of project values is sufficiently heavy-tailed, as for early-stage investments, a champions rule is optimal, while more consensus is optimal otherwise.

*Malenko: malenkoa@bc.edu; Nanda: ramana.nanda@imperial.ac.uk; Rhodes-Kropf: mat-trk@mit.edu; Sundaresan: s.sundaresan@imperial.ac.uk. Rhodes-Kropf is an advisor to, and investor in, Correlation Ventures, which provided some of the data. We are grateful to Rustam Abuzov, Sugato Bhattacharyya, Patrick Bolton, Felipe Csaszar, Andras Danis, Will Gornall, Jerry Green, Rustam Ibragimov, Steven Kaplan, Ali Lazrak, Dan Luo, Nadya Malenko, Uday Rajan, Morten Sorensen, Huseyin Yildirim, and Emmanuel Yimfor for helpful discussions as well as seminar and conference participants at the Financial Organisation and Markets Conference, LSE Finance, Michigan Ross, NBER Summer Institute, Vienna Festival of Finance Theory, University of Southampton, Utah Winter Finance Conference, and Western Finance Association. This paper is part of a project that has received funding from the European Research Council (ERC) Horizon 2020 research and innovation programme, Grant Number 865127. All errors are our own.

1 Introduction

Corporate decisions, whether they are about hiring, R&D financing or investment, are often made by a committee. When the outcome of those decisions is ex-ante uncertain, the goal of the committee is to effectively aggregate information into a decision with a positive expected present value. Theoretical work on information aggregation in voting, going back to the famous Condorcet’s jury theorem, has provided compelling arguments for the benefits of the ‘majority voting rule’ (Condorcet (1785); Ladha (1992); Feddersen and Pesendorfer (1997), Feddersen and Pesendorfer (1998)), or more generally, requiring some consensus to change the status quo decision. Using venture capital committee investing as a laboratory, this paper places bounds on the benefits of consensus in committee decisions. VC investments have substantial implications for economic growth¹ and as such, understanding their decision-making process is of first-order importance.

Using new evidence from a survey of the 50 largest U.S. venture capital (VC) investors, we show that while all use an investment committee to make decisions, a majority of these VC firms use a voting rule (either formal or informal) where the investment committee agrees to undertake the investment as long as at least one partner champions the deal (even if others are not bullish on the investment). This ‘champions voting rule’ is used primarily in seed and early stage investments, while in later stage investments, VC firms tend to move to more consensus-based voting rules. Even if the goal in seed and early stage investing is to ‘move quickly’ or ‘spray and pray’, it is significant that the selection rule for investments does not prioritise options where multiple partners’ quick reaction was positive, and instead allows for emphatic support by just one partner.

In a followup survey we sought to understand the reasons for the use of the champi-

¹As Lerner and Nanda (2020) show, VC-backed firms account for 89 percent of recorded R&D, and almost half of all non-financial IPOs in the past 20 years.

ons rule in early-stage investments. We offered three possible explanations for the use of champions voting rule for seed and early-stage investments but not for later-stage investments. These explanations are motivated by theory and were suggested to us by either academics or VC practitioners. The first explanation is that the champions voting rule is used because it is better at ‘catching outliers’ than the alternative decision rules. The second explanation is that early-stage deals require more specialized expertise that only few partners possess, so it makes sense to listen to whether the person with the most expertise in a particular project type champions for the deal or not. Finally, the third explanation is related to effort: earlier-stage deals are smaller in the amount of investment, so the partnership may want to have only one or few partners spend effort on diligence; in addition, the voting champion ‘owns’ the decision and thus has stronger incentives to ensure its success. We asked VC firms to evaluate the importance of each explanation. While all three explanations received some support, the overwhelming response was that firms use the rule to ‘catch outliers’, stating that the best early-stage investments are outstanding on some but flawed on many other dimensions and requiring more consensus risks missing the best investments as many partners would focus on the flaws. This evidence flies directly in the face of conventional wisdom, where voting is often considered valuable as a way to ‘trim outliers’ and emphasize the wisdom of the majority. These surveys raise a natural question: under what circumstances is too much consensus a bad thing?

To better understand the conditions under which the ‘champions voting rule’ arises as an optimal way to catch outliers and maximize expected returns for a committee, we evaluate our survey evidence through the lense of several canonical models of information aggregation. It is not a general implication of information aggregation models that the champions voting rule is optimal, even if catching right-tail outliers drives value. Thus,

the survey evidence is helpful at differentiating between different canonical models of information aggregation in committees.² We show that the survey evidence is inconsistent with models with binary project types as well as with models with infinitely many project types in which each committee member’s signal equals the true value of the project plus i.i.d. Normal error term, which has been a workhorse model of information aggregation in many settings.³ In particular, in this model the champions voting rule does similarly well at catching projects from the right tail as the majority voting rule, and even if the tail of the distribution of project values is very heavy, we do not find the champions voting rule to dominate the majority voting rule. Instead, we argue that the survey evidence is more consistent with the model in which the fundamental value of a project is determined by a number of dimensions, the quality of each dimension is independent from the other dimensions, and each partner gets an informative signal about one dimension but not the others.⁴

A key implication of this model is that the degree to which a voting rule with more or less consensus is optimal depends on the tail behavior of the distribution from which fundamentals are drawn. When the fundamentals observed by committee members are drawn from a subexponential – heavy right-tailed – distribution, the champions rule

²It is important to exercise caution in interpreting the survey results due to the potential disparity between the broad wording of the ‘catching outliers’ question and the specific dimensions of our proposed theory. Participants may have expressed agreement with the idea of ‘catching outliers’ based on their general understanding of its importance in VC investing, rather than explicitly aligning with the dimensions of our voting rule.

³This is a classic setting to study information aggregation in financial markets (e.g., [Hellwig \(1980\)](#), [Verrecchia \(1982\)](#), [Goldstein and Yang \(2017\)](#)). It has also been used to study information aggregation in committees ([Sah and Stiglitz \(1986\)](#), [Csaszar and Eggers \(2013\)](#)), and the optimal extent of experimentation in innovation ([Azevedo et al. \(2020\)](#)).

⁴In interpreting the conclusions drawn from our study, it is crucial to recognize that our model relies on the assumption that there exists information or knowledge among VC partners that cannot be shared or conveyed. This assumption underscores the importance of orthogonal knowledge in driving the voting process and capturing unique expertise. While we have simplified this representation by associating the knowledge with specific dimensions, it is worth noting that this knowledge can encompass a combination of expertise across various dimensions.

‘catches an outlier’ with a higher probability than any other voting rule, and when the distribution tails are sufficiently heavy, it emerges as the optimal voting rule to maximize expected returns. In contrast, when fundamentals are distributed with thinner right tails, majority voting rules tend to perform better.

The intuition here stems from two key attributes of heavy right-tailed distributions: first, the fact that the distribution of project payoffs exhibits heavy right tails implies that much of the potential value from investment comes from maximizing the probability of finding a superstar project. Indeed, as Bill Gurley, one of the leading VC investors, highlighted⁵:

... obviously you can't do every deal, you can't do every investment, you go broke. But getting overly jazzed about correctly identifying a negative or a no, it's just not that big a deal. It's not the job. The job is to find the outliers.

Second, voting rules that require many affirmative votes (such as majority or unanimous agreement) result in the undertaking of projects that tend to be good on many dimensions but are likely to miss those that are *exceptional* on few dimensions but are mediocre elsewhere. An important property of subexponential distributions – known as the ‘catastrophe principle’ implies that *a superstar venture is more likely to be driven by one exceptional characteristic and many mediocre ones rather than by many very good characteristics*. Together these two effects – the importance of getting decisions on superstar projects correct and the fact that a superstar project is likely to be driven by one superstar characteristic rather than many good ones – implies that the champions rule dominates other voting rules in contexts where the distribution of information and returns have heavy right tails. Other decision rules, such a majority rule or unanimous agreement would vote down superstar projects with a high positive probability – because it is possible and even likely that they have many weak characteristics.

⁵<https://tim.blog/2023/01/25/bill-gurley-transcript/>

In Appendix D, we provide quotes from the VC investors we surveyed suggestive of these mechanisms. Here we note two by Marc Andreessen, co-founder of VC firm Andreessen Horowitz, which show how VCs think about early-stage investments⁶:

Google, Facebook, eBay and Oracle all had massive flaws as early-stage ventures, but they also had overpowering strengths.

and

Aggregate scores [from all partners] don't correlate strongly with ultimate returns. With that approach, you get the mush in the middle, with no big flaws but no great strengths.

For VCs, significant flaws are not deal-breakers, as strong negative components in valuation can be compensated for by even stronger positives. Therefore, requiring consensus when every partner cannot see all the positives or negatives themselves, can result in rejecting the best deals.

Having discovered that the benefits from using a champions voting rule are related to the tail behavior of the fundamentals, we next turn to analyzing whether this relation is consistent with the results of our surveys. While finding the fundamental distributions is out of reach, we can find the ex-post realizations of investments by round. We use novel data on venture capital return multiples at the level of each startup's round of financing to calculate the distribution of return multiples for different investment stages. We are therefore able to directly corroborate the intuition that returns of late stage investments are substantially less skewed than seed and early-stage investments. Our estimates of the tail index show that the distribution of returns of seed-round investing are significantly more heavy-tailed than those of any of the other rounds, while the returns of late-round investing are significantly less so.⁷ To the extent that the *ex-ante* valuations for these

⁶Eisenmann and Kind (2014), page 8.

⁷We use Hill's estimator of tail indices with 20%, 10%, and 5% cutoffs, as well as a log-log rank size regression model. A more complete description is presented in Section 4.

stages are distributed in a manner similar to the *ex-post* returns, this would provide a rationale for why venture capital investors use the champions rule primarily in the seed and early-stage investments and migrate to more traditional voting rules for later stage investments. This is consistent with the literature on the topic: as documented by [Hall and Woodward \(2010\)](#) and [Kerr, Nanda, and Rhodes-Kropf \(2014\)](#), over half of startups receiving VC investment fail completely while a few generate enormous returns. [Scherer and Harhoff \(2000\)](#) provide evidence that returns in venture capital, and for technological innovation more generally have ‘heavy right tails’ and there have been explicit suggestions among practitioners that early-stage VC returns follow a power law (Pareto) distribution ([Thiel and Masters 2014](#)).

Lastly, we evaluate the quantitative importance of using a champions rule instead of traditional voting in contexts such as seed and early-stage venture capital, by analyzing a numerical example of the model whose inputs fit the VC context. We consider a fund with 5 partners that makes 25 investments and assume that the distribution of fundamentals has a similar tail index to the estimate of the tail index for seed investments. In our numerical example, the model-implied probability that the VC will be able to “catch a unicorn” (specifically, have at least one of the investments deliver a multiple that is 10X or more) is 3.2 times higher for the champions rule than a majority rule.

The theory provides an explanation for the use of a champions rule. And our data on returns provides support for the underlying assumptions of the model. However, it is possible that the correlations in the data could come from other reasons. While we have no way to rule out every other possible explanation, the follow-up survey mentioned earlier asked the VC firms why they use the champions rule. The overwhelming response was that firms use the rule to ‘catch outliers’, stating that the best early-stage investments are outstanding on some but flawed on many other dimensions and requiring more consensus

risks missing the best investments as many partners would focus on the flaws. This evidence supports our theory and flies directly in the face of conventional wisdom, where voting is often considered valuable as a way to ‘trim outliers’ and emphasize the wisdom of the majority.

The lower chance of selecting exceptional ideas in early-stage settings when the champions rule is not used is useful to put into context, because the dominance of the champions rule for subexponential distributions relies on the assumption that committee members have the same objective. If individual committee members have private benefits (e.g. [Scharfstein and Stein \(2000\)](#)), the champions rule can result in strategic championing for projects that are bad but have high private benefits. The cost of ‘over championing’ poor projects can over-ride the potential benefits from selecting outlier projects. In such an instance, more traditional voting rules are likely to be second best alternative, providing a potential rationale for the limited observed use of this type of voting – which we document for venture capital but does not appear to be widely used when selecting between potential projects within corporate R&D of large companies. Additionally, while there might be other factors affecting VC voting, such as disagreement, non-rational expectations, or inefficiencies in decision-making, we believe the simplified setting we consider captures the intuition underlying the advantage of the champions rule at “catching outliers.” However, we do consider a version of the model with behavioral biases in the appendix.

Literature Review

The paper is primarily related to a large literature on decision making in committees whose members have dispersed information, surveyed in [Gerling et al. \(2005\)](#). Most of the literature on committee decision-making and voting focus attention on the setting with binary states and signals. [Sah and Stiglitz \(1988\)](#) present a theory of optimal committee

voting rules in this environment, building on their earlier insights in [Sah and Stiglitz \(1986\)](#). In this framework, the problem of finding the optimal voting rule reduces to a statistical problem of finding the number of positive signals at which the committee is just indifferent between investing and not investing. While this setting can generate the optimality of requiring no consensus, it does not capture the intuition that the champions rule is better at ‘catching outliers’, while the majority rule results in ‘mush in the middle’ projects.

Two other strands of the literature on committee decision-making are related. First, several papers analyze models with the same information structure that we find most consistent with our survey evidence: the value of a project is a linear function of multiple independent signals, and each committee member learns one of them ([Moldovanu and Shi \(2013\)](#); [Malenko \(2014\)](#); [Name-Correa and Yildirim \(2019\)](#)). In particular, Section 5.3 of [Name-Correa and Yildirim \(2019\)](#) studies optimal voting rules in this model when the distribution of signals is an interval and the uninformed decision-maker is indifferent between the two actions. The element that we focus on is the relation between the tail behavior of the state and signal distributions and the optimal voting rule.

Second, several existing papers establish the optimality of the unanimity rule in various environments ([Coughlan \(2000\)](#); [Bond and Eraslan \(2010\)](#); [Jackson and Tan \(2013\)](#); [Chan et al. \(2018\)](#)). In [Coughlan \(2000\)](#), unanimity can be optimal if committee members can communicate all information prior to the vote and they have similar preferences. In [Bond and Eraslan \(2010\)](#), the unanimity rule is beneficial because it incentivizes the proposer to make a proposal that is attractive to the rest of the group. In [Jackson and Tan \(2013\)](#), the unanimity rule is beneficial because it encourages committee members to disclose verifiable information prior to the vote. Finally, in [Chan et al. \(2018\)](#), the unanimity rule can be beneficial because in an environment with committee members

with heterogeneous discount factors unanimity makes makes patient members pivotal, leading to more information acquisition and more precise decisions. However, since the unanimity rule is the opposite of the champions rule, the forces highlighted in these papers work against explaining the use of the champions rule by VC investment committees.

Our work also builds on a long literature that has examined venture capital’s role in financing innovation. In particular, this research has fleshed out many of the tools venture capital investors use to improve the outcome of the startups they back, such as staged financing ([Gompers 1995](#); [Bergemann and Hege 1998](#)), securities that have state-contingent cash flow and control rights ([Hellmann 1998](#); [Cornelli and Yosha 2003](#); [Kaplan and Strömberg 2003](#)) and the active role of venture capital investors on boards of portfolio companies ([Hellmann and Puri 2000, 2002](#); [Lerner 1995](#)). Our work builds on the nascent literature on understanding decision making in venture capital partnerships including how venture capital investors select investments ([Kaplan, Sensoy, and Stromberg 2009](#); [Gompers et al. 2020](#)), which is an important element to understanding the role of venture capital investors in financing innovation ([Lerner and Nanda 2020](#)). This work is related to the broader literature on the incentive, agency and organizational frictions among intermediaries financing innovation ([Manso 2011](#)), the role this can play in surfacing transformational ideas ([Bloom et al. 2020](#)) and the fact that radical innovations that upend existing firms often arise from venture capital despite the much larger R&D expenditure directed towards the financing of innovation in large companies across the world ([Kortum and Lerner 2000](#)).

The remainder of the paper is organized as follows: Section 2 provides survey evidence on voting practices and justifications for VCs. Section 3 lays out a model of voting and shows conditions for the optimality of a champions voting rule. Section 4 shows empirical evidence consistent with the conditions from Section 3 and provides a quantitative exercise

to measure the size of the benefits. Section 5 concludes.

2 Survey Evidence on VC Voting Rules

Venture capital provides a unique context within which to examine empirical patterns related to committee decision making for several reasons. First, VC partnerships view project selection as among the most important determinants of their success (Gompers et al. (2020)) and in addition appear to exhibit substantial heterogeneity in the ways in which they make investment decisions. Second, it has been well-documented that venture capital returns are driven by a few outliers (e.g., Hall and Woodward (2010) and Scherer and Harhoff (2000)), often referred to as ‘home runs’ by practitioners. Indeed, there has even been explicit suggestions among practitioners that early-stage VC returns follow a power law (Thiel and Masters (2014)). We conducted two surveys, the key questions of which can be seen in Appendix D.

2.1 Evidence on Voting Rules

We follow the examples of Graham and Harvey (2001) and Gompers et al. (2020) to survey VC investors on their voting practices. While our empirical approach builds on and is most similar to Gompers et al. (2020), we note some key differences. Gompers et al. (2020) have an extensive survey of over 650 VC investors across a wide range of topics. Our approach focuses in more detail on voting practices within investment committees, and also aims to look at those who invest across multiple stages so as to get within-VC variation in voting practices across rounds. Since it is only larger VC funds that typically have the ability to invest across seed, early and late stage, we focus our survey on U.S based managing partners at 55 the largest VC firms that make investments into U.S

startups. Our measure of size is based on the cumulative fund raising over the 2016-2018 period as calculated from Pitchbook. Figure 12 documents the key questions used for this analysis as they were posed in the survey. As can be seen from the questions, we asked about both the formal voting process used in these VC firms as well as the informal process. We received responses from 35 of these 55 firms, implying a response rate of nearly two-thirds.

As noted above, the VC firms we targeted were larger investors. Despite our focus on a narrow sample of VCs, it is important to recognize that these investors are responsible for a disproportionate share of the dollars invested into VC-backed startups in U.S. and hence are more representative of VC investing than might be expected. For example, [Lerner and Nanda \(2020\)](#) examine fund raising by VC investors between 2014 and 2018 and find that the top 50 VCs (or approximately 5% of those who raised funds in that period) accounted for half of the total capital raised over that period.

The average investment committee at the VC firms we surveyed had 10.4 partners compared to an average of 4.8 partners in the VC firms surveyed by [Gompers et al. \(2020\)](#). Despite this difference and the fact that over half our sample comprised VCs investing from funds over \$500 million, the partnership size in VCs remained relatively small. VC firms in the 75th percentile in our sample had 13 partners on their investment committee. In other words, the size of investment committees in VC firms appears not to scale proportionately to the size of the funds. This lack of scaling among the partners on investment committees is an interesting fact, one that seems different from partners at other professional service firms such as lawyers and management consultants.

Having documented the characteristics of the VC firms and the questions asked of the respondents, we turn next to outlining the key results. The charts report results broken down by the stage of the investment being considered when the partners are voting. The

first bar corresponds to “Seed” stage investments, which are the earliest investments into startups and are believed to have the most skewed returns. The next bar corresponds to “Early Stage” investments, typically considered to be Series A and at times, Series B investments. The final bar corresponds to later stage or “Growth Stage” investments, which are typically made into more mature startups which have already shown some degree of product market fit, are often generating some revenue and hence are the least skewed in terms of the profile of returns.⁸

Within each of these stages, we further break down the results by share of the firms that report using different types of voting rules. As can be seen from the bars in Figure 1, 60% of all VCs in our sample decide whether to deploy capital into a Seed stage startup by using a champions voting rule in their investment committee, where a single partner can go ahead and do the deal regardless of what the others feel. A further 30% of VCs use a variant of the champions rule, where a single champion can do the deal as long as there is no veto. For investments into early-stage ventures, the share of VCs using champions voting falls to 20% and a much larger proportion of VCs use some form of majority or consensus to decide which investment to make. By the growth stage, this share falls even further. It is worth emphasizing that since most of these VCs invest across all stages, a shift in the share of VCs voting using the champions rule across the different stages is evidence of the same VCs changing their voting rule across stages. This provides compelling evidence of voting rules shifting by stage, as the variation being documented is “within VC” as opposed to “across VCs”, the latter of which is much more subject to concerns about unobserved heterogeneity.

In Figure 2, we further document that this pattern continues to exist in a very similar manner when one examines the informal voting practices within the investment commit-

⁸We use novel data to validate the difference in skew across rounds of financing in Section 4.

tees. At the Seed stage, an enthusiastic champion is sufficient for investment in 90% of investment committees regardless of the formal voting rule in place, implying that in practice the veto is rarely used. However, this deference to the enthusiastic champion on the investment committee falls within the *same VC firm* by the Early stage, with a greater proportion of investment committees requiring either a majority of individuals to be enthusiastic, or all individuals to be enthusiastic. By the growth stage, the informal process reflects even less deference to an enthusiastic champion.

2.2 Evidence on Voting Reasons

We turn next to understanding why these different voting practices might be undertaken at VC firms, in particular the emphasis on champions voting when investing in extremely early stage startups.

The same set of 55 firms were sampled, and we received 19 responses, implying a response rate of a little over a third. The average investment committee at the responding firms had 12.8 partners. The most recent fund raised for 16 of the 19 respondents was at least \$500 million, while the remaining three were between \$150 million and \$500 million.

We told the participants the results of our initial survey: that champions voting appears to be prevalent on early-stage deals, and proposed three explanations, suggested to us by colleagues that we shared the results of our initial survey with. The first explanation was that “earlier-stage deals require[d] more specialized expertise than late-stage deals, so partnerships give people with expertise greater discretion in doing earlier-stage deals but not later-stage deals.” Formally, only one or few partners get informative signals about early-stage deals, but more partners get informative signals about later-stage deals. The second explanation was that “the best early-stage investments can be extremely promising in some dimensions but flawed in others. Requiring too much consensus can therefore

lead partnerships to pass on the best deals.” The formal model that shows this logic is presented in the following section. The last explanation was that “VCs care less about smaller deals, and don’t want too many partners to spend effort on diligence. Leaving the decision to a champion for these smaller deals also incentivizes greater effort on the part of the champion.” Formally, each partner needs to incur an effort cost to get an informative signal about the project. If there is a fixed cost of effort, it can be optimal to have only one partner acquire information.⁹ If degree of signal precision is endogenous and one very precise signal is better than several moderately precise, then concentrating information production on one partner can result in a better decision.¹⁰ For each explanation, the participants were asked if they “strongly agreed”, “somewhat agreed”, “neither agreed nor disagreed”, “somewhat disagreed” or “strongly disagreed”.

We present the key results in two figures. In the first, Figure 3, we show the percentage of respondents who agreed with the three explanations we offered. The first set of three bars shows the percentage who either “somewhat agreed” or “strongly agreed”, while the second set shows the percentage who “strongly agreed”. The dark blue bars correspond to the “catching outliers” explanation; the light blue bars correspond to the “expertise” explanation, while the grey bars correspond to the “effort” explanation. While all three enjoy at least some support, that support is clearly greatest for the “catching outliers” explanation, with over half (58%) of respondents strongly agreeing with it, and 84% at least somewhat agreeing.

The second, Figure 4 shows the percentage of the time one explanation was ranked more highly by respondents than others. The first set of three bars shows the fraction

⁹See [Yung \(2005\)](#) for a model based on this logic that shows that underwriters benefit from restricting participation in an IPO.

¹⁰See [Khanna and Mathews \(2011\)](#) for a model in which herding on an initial decision in a multistage context can result in more informative and better outcomes due to a related intuition.

of the responses that had each explanation as the joint or outright leader; the second set shows the fraction of responses that had each explanation as the outright leader; the third set shows the fraction of responses that had each explanation dominated by at least one other. Again, “catching outliers” is most frequently the favorite explanation from respondents, dominating 84% of the time, while it is only dominated by another explanation in 16% of responses.

Given that VCs tend towards champions voting, and that their justification for doing so explicitly refers to catching outliers, we next set up and solve a model to find the conditions under which doing so would maximize expected value.

3 Interpreting Survey Evidence via Models of Information Aggregation in Committees

To sum up, the evidence from our surveys and the stated reasons why VC investment committees use certain voting rules can be summarized in the following way:

- Investment committees use the champions rule for seed and sometimes for early-stage investment proposals. The same investment committees tend to shift to the majority rule when they evaluate potential investment in later-stage projects.
- When asked why they adopt the champions voting rule, the explanation that receives the most support is that the champions voting rule performs better at “catching outliers” than other voting rules.
- There is belief that requiring a lot of consensus results in “mush in the middle” investment decisions, with no big flaws but no great strenghts.

Our next goal is understand which models of committee decision-making are consistent with these facts. While “effort” and “expertise” explanations received some support in our second survey, they are not the explanation that received the most support. As a consequence, we abstract away from the issues of effort in information acquisition and from asymmetric expertise among committee members and focus on purely information aggregation models with symmetric committee members.

Specifically, suppose that there are N partners at a venture capital firm, who need to decide whether to accept ($a = 1$) or reject ($a = 0$) a project. The project has an upfront investment cost $I > 0$ and yields a payoff V upon success. Each partner gets a signal θ_i about payoff V . We consider voting mechanisms where every partner simultaneously submits a binary vote $v_i \in \{0, 1\}$. The project is deemed worthy of investment if and only if the total number of votes exceeds some cut-off k , i.e., $\sum_{i=1}^N v_i \geq k$. This voting mechanism captures as special cases the following three decision-making rules:

1. Champions rule ($k = 1$): The fund undertakes the investment if and only if there is at least one partner that votes (“champions”) for it.
2. Simple majority rule ($k = \frac{N+1}{2}$): The fund undertakes the investment if and only if $\frac{N+1}{2}$ or more partners vote for it.
3. Unanimity rule ($k = N$): The fund undertakes the investment if and only if no partner objects to it.

A natural question in this voting model (and many others) is why agents do not just announce their information to the group, thus always achieving the best outcomes. As we show in two extensions at the end of this section, communication, even noisily can weaken the results of the model, but in the absence of perfect communication, the paper’s results

will hold. We assume for all of our results that perfect communication of a partner’s private information is impossible (or alternatively, so prohibitively costly as to never be optimal).

This setup captures a general model of information aggregation in committees. The specific models differ in their assumptions about the joint distribution of project payoffs and signals $(V, \theta_1, \dots, \theta_N)$. Ideally, one would estimate this joint distribution from the data. This, however, is not feasible. Instead, our approach is to analyze the empirical evidence through the lense of several canonical models of information aggregation, and conclude which ones of them, if any, is consistent with the empirical evidence.

3.1 Models with Binary Signals

Models with binary signals are probably the most common models used in the literature on information aggregation in voting and committee decision making.

A natural starting point to think about optimal voting rules in committees is [Sah and Stiglitz \(1988\)](#), who studied optimal voting rules in a committee when there are two kinds of projects (good and bad) and each committee member gets a binary signal about the project’s type, independent conditionally on the project’s type. In this case, the search for the optimal voting rule is equivalent to a statistical problem of finding the ratio of positive to negative signals at which the committee is indifferent between investing and not investing. The optimal voting rule is determined by the fraction of good projects in the pool and costs of Type I and Type II errors. The champions rule is optimal if the cost of Type-II error is sufficiently higher than the cost of Type-I error and the prior probability that the project is good is sufficiently high. In the application to early-stage VC investments, the first condition is very likely to be satisfied: The cost of missing on a good investment is much higher than the cost of making a bad investment. In contrast,

the second condition is likely to be violated: VCs invest in a very small fraction of early-stage investments they evaluate, so the prior probability that the project is good is very low. Nevertheless, if the first force is sufficiently powerful, the champions rule can be optimal in this setting.

However, a model with two project types cannot distinguish between good projects and outstanding projects, so there are no notions of “passing on the best deals” and “the mush in the middle.” So overall we conclude that these models do not do a very good job at capturing the empirical evidence on the behavior of VC investment committees.

3.2 Model with Signals Equal to State Plus i.i.d. Noise

We next consider a model in which each committee member i receives a signal that equals the true value of the project with i.i.d. noise:

$$\theta_i = V + \varepsilon_i,$$

where ε_i is an i.i.d. draw from some distribution $F(\cdot)$. When distribution of noise is Normal, this setup captures a workhorse model of information aggregation in financial markets (e.g., [Hellwig \(1980\)](#), [Verrecchia \(1982\)](#), [Goldstein and Yang \(2017\)](#)). In the case of aggregation of information in committees, this setup has been used in [Sah and Stiglitz \(1986\)](#) and [Csaszar and Eggers \(2013\)](#), among others. In the context of innovation, [Azevedo et al. \(2020\)](#) use this additive setup with normally distributed errors and value being drawn from a distribution with fat tails to study the optimal extent of experimentation. In equilibrium, each committee member i votes for the project if and only if

her signal θ_i exceeds the threshold level θ_k^* , determined by

$$\mathbb{E}[V|\theta_1 = \theta_k^*, \theta_2, \dots, \theta_k \geq \theta_k^*, \theta_{k+1}, \dots, \theta_N \leq \theta_k^*] = I. \quad (1)$$

Intuitively, a committee member i knows her signal θ_i and anticipates that her vote only matters when the votes of others are split: $k - 1$ other committee members vote for the project and $N - k$ committee members vote against the project. This event implies that $k - 1$ signals are above the equilibrium threshold θ_k^* and $N - k$ signals are below θ_k^* . Condition (1) means that the committee member with signal θ_k^* is just indifferent between investing and not investing given this information.

Since this model does not admit analytic solutions, we analyze whether this model can explain the empirical evidence on voting in venture capital committees by simulations. Specifically, we simulate draws of V from Pareto distribution with various degrees of scale and shape parameters. Our baseline case has $N = 5$ committee members, shape $\alpha = 1.7$ (corresponding to the empirical estimate in Section 4), the investment cost $I = 1$, the scale parameter $x_m = 0.25$, and the standard deviation of signal error of one, but we repeat the analysis for many other parameter values. After numerically calculating the equilibrium threshold θ_k^* for each decision rule k , we perform three analyses. First, we calculate the optimal decision rule, i.e., find the value of k at which the expected value from the investment decision is maximized. We do this for a range of different parameters. Second, for each decision rule k , we calculate the distribution of payoffs of accepted projects that are implied by each decision rule. Finally, for each decision rule we calculate the probability that a project with value V gets accepted, $p_k(V)$. We find that the champions decision-making rule is not optimal regardless of shape α of distribution of V .

For our baseline parameters, the optimal decision rule is simple majority ($k = 3$). In

fact, keeping the other parameters constant, the majority rule is optimal for any shape α of distribution V . When the investment cost I is sufficiently high (or, equivalently, scale x_m is sufficiently low), we obtain that the optimal decision rule requires $k = 4$ positive votes out of $N = 5$.

Further, we find that the decision-making rule has little effect on the investment in superstar (very high V) projects and primarily changes investment decisions in marginal projects.

This effect can be seen in Table 5 and Figures 8 and 9. Figure 9 plots the distribution of values of accepted projects for the champions, majority, and unanimity decision-making rules, respectively. As one can see, different decision rules imply different return distributions on accepted projects that the partnership is getting. However, these differences do not occur for projects in the right tail. Intuitively, conditional on a project having extremely high fundamental value, it is very likely that each partner's signal will also be high, and so the project would receive support in any decision rule. The differences are pronounced among low-quality ($V < 1$) and medium quality ($V \in (1, 2)$) projects, where the majority rule achieves a better distribution of project returns (more high-quality and fewer low-quality projects) than the champions rule. Table 5 corroborates this evidence with quantile statistics. In particular, it shows that the right tail behavior of project payoffs is quite similar across the decision rules, but quite different for projects in the bottom 50% of values.

Figure 8 plots probabilities of a project with value V getting accepted for three decision rules for different values of V . It shows that projects from the right tail of the distribution of values get accepted with probability close to one regardless of the decision rule. Intuitively, for any distribution of noise, if V is sufficiently high, the probability that signal $V + \varepsilon_i$ exceeds the equilibrium voting cutoff is close to one. However, the

majority rule is significantly more efficient at selecting moderately good projects (e.g., $V \in (1, 2)$) than either the champions or the unanimity rules. This result captures a common intuition that the majority rule is good at averaging out the noise and trimming outliers.

Overall, this model illustrates the intuition that the majority rule is good at trimming outliers. However, the implications of this model are clearly inconsistent with our empirical evidence on decision-making in VC committees.

3.3 Model where State Equals the Sum of Characteristics

Finally, we consider another canonical model from the literature on decision-making in committees. In this model, the project is determined by the values of a set of M -many characteristics where $M \geq N$:

$$V = \theta_1 + \theta_2 + \dots + \theta_M. \tag{2}$$

All M characteristics are distributed i.i.d. over $[-l, \infty)$ (for some constant $l \in \mathbb{R}$) according to a distribution function $F(\cdot)$ with density $f(\cdot)$. Let $G(\cdot)$ denote the implied distribution of V . Each partner i perfectly observes the value of characteristic i . Thus, partner i learns θ_i but only knows the distribution of the other characteristics. This model has been quite popular in the literature on decision-making in committees and organizations more generally, in particular, because it leads to tractable solutions. Variations of this model have been studied by [Moldovanu and Shi \(2013\)](#), [Malenko \(2014\)](#), [Name-Correa and Yildirim \(2019\)](#), [Harris and Raviv \(2005\)](#), and [Harris and Raviv \(2008\)](#). As we will argue below, this model is most consistent with our empirical evidence on decision-making in VC committees.

Independence of characteristics is a strong assumption, but is used in this case to highlight the value of additional information that each partner brings; also note that as characteristics become perfectly correlated then voting is no longer needed. One can more generally interpret θ_i as the portion of partner i 's information about residual uncertainty of the project's value. For example, if the value of each characteristic were a sum of the common factor Z and an idiosyncratic factor and each partner learned the values of both, the model becomes similar, as we can simply subtract the common factor from the investment cost. Intuitively, the idea behind (2) is that different partners at a firm might have different, and orthogonal areas of expertise in assessing a company's value, and will be better placed to assess the values of those characteristics. For example, one partner can be an expert in assessing the technology, another partner can be an expert in assessing potential demand for the product, while the third partner can be an expert in assessing the quality of managerial team of the start-up.

The assumption in our model that each VC partner focuses on a single dimension was made for the purpose of providing clarity and intuition. In reality, this simplification does not imply that partners exclusively consider a single dimension. Rather, it signifies the presence of orthogonal knowledge that is not fully conveyed. This remaining orthogonal knowledge represents unique insights and expertise that contribute to the voting process. Future research could explore more nuanced decision-making frameworks that incorporate multiple dimensions of knowledge within the VC partnership.

We assume that $\mathbb{E}[V] \leq I$, which is equivalent to $ME[\theta] \leq I$. In the context of early-stage investment, it is natural to model F as having heavy tails. As we show in the next section, this assumption fits the empirical evidence: heavy tails of the characteristic distributions imply in the model heavy tails of the return distribution, which is strongly supported by the data, especially for early-stage investments. Further, venture capital

firms find it very important to “catch the unicorn” (i.e., find and invest in projects with very high V s). In contrast, in later stage investments, it is natural to expect that the distribution of valuations has thinner tails: a project is unlikely to be superstar if it does not already have high profile by then.

3.3.1 Optimality of the Champions Rule when Tails are Sufficiently Heavy

We next solve for the optimal voting rule k . In particular, we are interested in the conditions under which the champions rule ($k = 1$) arises as the optimal one.

Consider a voting rule that requires k positive votes for approval of the project. Under this rule, each partner will vote for the project if and only if her characteristic’s value exceeds some cut-off $\hat{\theta}_k$. Threshold $\hat{\theta}_k$ is the value such that each partner is exactly indifferent between investing and not investing, given her characteristic and the fact that her vote is pivotal:

$$\hat{\theta}_k + (k - 1) \mathbb{E} [\theta | \theta \geq \hat{\theta}_k] + (N - k) \mathbb{E} [\theta | \theta \leq \hat{\theta}_k] + (M - N) \mathbb{E} [\theta] = I. \quad (3)$$

Equation (3) pins down the voting threshold $\hat{\theta}_k$. It satisfies the intuitive property that $\hat{\theta}_k$ is decreasing in k . This follows from monotonicity of $\mathbb{E} [\theta | \theta \geq \hat{\theta}]$ and $\mathbb{E} [\theta | \theta \leq \hat{\theta}]$ in $\hat{\theta}$. Intuitively, each partner is more aggressive about voting for the project if more votes are needed to approve the project.

From (3), it is easy to see the pros and cons of the champions rule and more conventional voting rules, such as simple majority. On the one hand, a good project but with only few “superstar” characteristics will be rejected by the majority rule, if its other characteristics are weak. In contrast, the committee will invest in this project under the champions rule, because the partner with a very strong characteristic will champion for

it. On the other hand, a good project with many good characteristics will not be invested in under the champions rule, if none of its characteristics are of superstar quality. Since these two types of projects potentially result in very different return profiles, the choice of the information aggregation rule is akin to the choice of the return profile implied by each rule.

Our first proposition shows that the champions rule does particularly well in picking “superstar” projects if the distribution of characteristics has sufficiently heavy tails (formally, it is subexponential):

Proposition 1. *Suppose that F is subexponential, and let $d_k(\theta) \in \{0, 1\}$ denote the investment decision under voting rule k when the realized vector of characteristics is θ . Consider any voting rule $k > 1$. If $\Pr(d_k(\theta) = 1) - \Pr(d_1(\theta) = 1)$ is either non-positive or positive but not too high, then $\lim_{\hat{V} \rightarrow \infty} \Pr(d_k(\theta) = 1 | V > \hat{V}) < \lim_{\hat{V} \rightarrow \infty} \Pr(d_1(\theta) = 1 | V > \hat{V})$, that is, the champions rule accepts a superstar project with a strictly higher probability.*

Subexponential distributions is a subclass of heavy-tailed distributions whose tails decrease slower than any exponential tail. Almost all commonly-used heavy-tailed distributions are subexponential: for example, Pareto (power law), Weibull (with $\alpha < 1$), and lognormal distributions. This class of distributions is used frequently in the analysis of insurance claims and rare events. The reason for Proposition 1 comes from an important property of subexponential distributions called the “catastrophe principle.” The catastrophe principle says that the distribution of a sum of N subexponential random variables in the tail is similar to the distribution of the maximum element in the sum. Informally, it means that a superstar project is much more likely to be driven by one superstar characteristic and many mediocre ones rather than by all very high characteristics. This property is consistent with the fact that Facebook, Google, eBay, and other super-

star companies all had many flaws as early-stage ventures but also some overpowering strengths, as Marc Andreessen’s quote above asserted.

The intuition for Proposition 1 is that if a superstar project is driven by one superstar characteristic and one of the partners gets an informative signal about it, the champions rule would accept this project, since the informed partner would champion for it. In contrast, any other voting rule would reject this project with positive probability, which happens if the other $N - 1$ characteristics that partners get signals about are relatively weak. The role of the condition on the unconditional probability of investment in the proposition is the following. If the characteristic that drives the superstar project is not learned by any of the partners, then the probability of catching this project is given by the average probability of investment. As long as it is either lower under voting rule k than under the champions rule or higher but not too much higher, the effect that the probability of investment is highest under the champions rule when one of the partners gets a signal about a superstar characteristic dominates.¹¹ Note that the fact that F is subexponential is an important condition for Proposition 1. Specifically, if F has light tails (e.g., Normal), a superstar project is typically driven by many characteristics that are pretty good rather than by one superstar characteristic. In this case, the advantage of the champions rule over other decision making rules in identifying superstar projects is lost.

While Proposition 1 implies that the champions rule is more likely to identify projects with superstar payoffs than any other voting rule if the distribution of characteristics is subexponential, it does not imply that it leads to a higher expected payoff because it may be more likely to miss non-superstar good projects, and a typical project is the latter.

¹¹In all our examples, the probability of investment is higher under the champions rule than under any other voting rule, so this condition is trivially satisfied.

However, as the next two propositions show, this will be the case if the right tail of the characteristic distribution¹² are sufficiently important.

Proposition 2. *Suppose that the right tail is sufficiently important in the sense that $\int_{\hat{\theta}_1}^{\infty} \theta dF(\theta)$ is sufficiently close to $E[\theta]$ and $1 - F(\hat{\theta}_1)$ is sufficiently close to zero. Then, the champions rule has a higher expected value than any voting rule $k > 1$.*

The intuition for Proposition 2 naturally follows from Proposition 1. Recall that Proposition 1 shows that when the distribution of characteristics has heavy tails, the champions rule is better at identifying superstar projects than any other decision making rule. Proposition 2 shows that if identifying projects in the tail is sufficiently important, then the advantage of the champions rule outweighs the advantage of other decision making rules.

While Proposition 2 helps at showing the intuition, one may wonder if it is relevant in realistic settings. The next result specializes to the case of Pareto (power law) distribution, which is a popular distribution to depict returns on venture capital investments, and obtains a cut-off $\hat{\alpha}$ on the shape parameter, such that the champions rule is optimal for any $\alpha \leq \hat{\alpha}$.

Proposition 3. *Suppose that θ s are distributed according to the Pareto Type 1 distribution with shape parameter α . For any fixed and finite mean level of the characteristic $E[\theta]$, there exists cut-off shape parameter $\hat{\alpha}$, such that the champions voting rule is optimal for any $\alpha \leq \hat{\alpha}$.*

The argument for Proposition 3 is related to Proposition 2, and the broad intuition remains the same. If the tails of the distribution are sufficiently heavy (i.e., α is sufficiently

¹²And consequently, of the return distribution, since for subexponential distributions, the tail distributions of the characteristic and the sum coincide.

low), then (1) correctly identifying projects with very high payoffs is a bigger driver of value than correctly identifying projects with moderately high payoffs; and (2) these projects are more likely to be driven by one very high realization of θ rather than by multiple moderately high ones.

3.3.2 Illustration and Comparisons

Figure 5 illustrates Proposition 3. Here, under the assumption that characteristics are distributed according to a Pareto distribution with shape parameter α , we plot the optimal value of k as a function of α and the number of investors N . It is clear to see that the champions rule (dark blue color) is often the optimal rule for even moderate values of α , but that, more pertinently, for every value of N , there is a threshold α below which champions rule is always optimal.

By analogy with Figures 8 and 9 and Table 5, we show the analogous statistics for the model in which the true value of the project equals the sum of independent project characteristics, and each partner learns one of them. Figure 11 plots the distribution of values of accepted projects for the additive model for the champions, majority, and unanimity decision-making rules, respectively. Similarly to Figure 9, different decision rules imply different return distributions on accepted projects. However, unlike in the case of Figure 9, now different decision rules imply very different return distributions in the right tail. Specifically, projects in the right tail are much more common for the champions rule than for majority and unanimity rules. Table 5 illustrates this point via quantile statistics. For example, the 99th percentile of the payoff of accepted projects (as a multiple of the investment cost) is 7.8 for the champions rule, but 3.6 for the majority rule and 1.6 for the unanimity rule. In contrast, projects around the zero-NPV threshold are more common for majority and unanimity decision rules than for the champions

rule. In this respect, the majority rule selects “the mush in the middle.” Figure 10 plots probabilities of a project with value V getting selected for the three decision rules. The champions rule is much more likely to accept any project, both good and bad, than other decision rules, but the difference is especially pronounced for projects with values in the right tail.¹³

Overall, this model is more consistent with our empirical evidence than the alternatives. In particular, it captures three motivating pieces of empirical evidence. First, the champions rule is optimal for seed projects but not for later-stage investments. This is because the distribution of payoffs (and, thus, signals) has significantly fatter tails for seed projects than for later-stage investments. Second, the reason why the champions rule is optimal in the model is consistent with the explanation provided by the VC committees that the champions rule performs better at “catching outliers.” Finally, this model predicts that decision rules that require more consensus result in “mush in the middle” projects.

Finally, it is worth contrasting the case in which superstar projects are driven by few superstar characteristics with the opposite case in which for a project to be successful, all its characteristics must be of sufficiently high quality. Specifically, suppose that the mapping of project characteristics into project value is not additive, as in (2), but rather $V = \min \{\theta_1, \dots, \theta_M\}$. In other words, the value of each project is determined by its worst characteristic. This specification can be motivated in the same way as the O-ring theory of economic development (Kremer (1993)): one bad component can make an otherwise successful project fail, as in the case of the O-ring imperfection leading to a collapse of the space shuttle Challenger. The next proposition shows that the consensus rule is optimal

¹³We expect the reality to be not as extreme as illustrated in Figure 10 for two reasons. First, in reality the values of different characteristics are likely positively correlated, rather than uncorrelated. Second, in reality some of the information gets aggregated in pre-vote communication.

for projects of this type:

Proposition 4. *Suppose that $V = \min \{\theta_1, \dots, \theta_M\}$. Then, regardless of the distribution of characteristics, the optimal voting rule is consensus ($k = N$).*

Intuitively, unlike any other decision rule, the consensus rule ensures that all characteristics are above a certain bar, which is very valuable when the value of the project is determined by its weakest characteristic. In fact, as we show in the proof, the consensus rule implements the same investment decision as the planner with knowledge of all M characteristics $\theta_1, \dots, \theta_M$. Thus, the consensus rule is optimal among not only all voting rules but also among all general mechanisms.

3.4 Extensions

Til now we have precluded partners from communicating with each other prior to voting. The assumption of imperfect communication is required to set up any voting model with common interests - if agents could perfectly communicate, there would be no need to vote, as they would always perfectly agree on the correct course of action. However, we can also accommodate partial communication in our framework. Specifically we consider two extensions to the model in which the value equals the sum of signals. In the first every partner i can noisily communicate their private θ_i to their colleagues. In the second, every partner has a positive but non-guaranteed probability of perfectly communicating their private θ_i to their colleagues.

3.4.1 Voting with Noisy Communication

Suppose that partners are allowed communicate with each other, but can only do so noisily. As before, assume that $V = \theta_1 + \dots + \theta_M$, and that agent i sees θ_i perfectly. For

simplicity, assume that there are as many agents as dimensions of valuation ($N = M$).¹⁴ As before, each θ_i is distributed independently over $[-l, \infty)$ with $l > 0$, with distribution function $F(\cdot)$ and density $f(\cdot)$. Each partner i communicates a message t_i about θ_i to all of her colleagues. Let $h(\theta_i|t_i)$ denote the posterior distribution of characteristic θ_i conditional on message t_i . Assume that likelihood ratio $\frac{h(\theta_i|t_i)}{h(\theta_i|t'_i)}$ is strictly increasing in θ_i for any t_i and $t'_i < t_i$, and normalize messages so that each message equals the posterior expected value of a characteristic: $t_i = \mathbb{E}[\theta_i|t_i]$. Then, we can write down the value of the project as

$$V = \sum_{i=1}^M t_i + \sum_{i=1}^M (\theta_i - t_i).$$

Put differently, we can reframe posterior expectations such that the common value to all voters is not the ex-ante expectation $\mathbb{E}[V]$, but the post-communication expectation $\mathbb{E}[V|t_1, \dots, t_M] = \sum_{i=1}^M t_i$. Then each agent i 's private information is not their signal θ_i , but the *uncommunicated* part of that signal $\epsilon_i = \theta_i - t_i$. In general, the optimal voting rule k will depend on the distribution of uncommunicated parts of the signal ϵ_i and on how much information was communicated ($\sum_{i=1}^M t_i$). Nevertheless, notice that if ϵ_i is distributed subexponentially and $\sum_{i=1}^M t_i \leq I$, then Propositions 2 and 3 apply to this model directly, i.e., the champions rule is optimal under similar conditions.

3.4.2 Voting with Imperfect Communication

Another way to think of communication is to instead assume that each i , with some probability p , is able to perfectly communicate the value of θ_i to her colleagues. With probability $1 - p$, partner i communicates nothing, and her colleagues maintain their prior beliefs about θ_i . Therefore, after communication of all agents, some subset $S \subset \{1, \dots, N\}$

¹⁴Relaxing this assumption will change nothing besides some notation.

of agents will have communicated their beliefs perfectly, (where $|S| = n$ is the number of partners that successfully communicated their signals) while the remaining subset $T = \{1, \dots, N\} - S$ have failed to communicate at all. Therefore, among N partners, each partner $i \in S$ has no private information at the voting stage, while each partner $i \in S$ privately knows θ_i .

We next argue that Proposition 1 extends to this model. Again, for simplicity we focus on the case $M = N$. Consider a project with $V > \hat{V}$ for the limit case $\hat{V} \rightarrow \infty$. Since $F(\cdot)$ is subexponential, the distribution of V in the right tail coincides with the distribution of its highest characteristic $\max\{\theta_1, \dots, \theta_n\}$. There are two possible cases. With probability p , the highest characteristic is communicated successfully prior to voting. With probability $1 - p$, it is not communicated prior to voting. In the former case, the project is accepted with certainty under any voting rule. In the latter case, the project is accepted with certainty under the champions rule ($k = 1$), but with probability strictly below one for any voting rule $k > 1$. This follows directly from the proof of Proposition 1 (under the assumption that $M = N$) and the fact that no signal is communicated with strictly positive probability. Thus, the result of Proposition 1 also holds in this model.

Similarly, if the communicated signals are such that the post-communication expected value of the project, $\sum_{i \in S} \theta_i + (M - n) \mathbb{E}[S]$, is weakly below its investment cost, then the arguments of Propositions 2 and 3 apply.

These two extensions should reassure the reader that communication does not negate the fundamental conclusions of our model. However, it remains the case that communication will typically tend to reduce the importance of any voting rule, as the problem of information aggregation in voting will be weakened.

3.4.3 Other Extensions

We have analyzed three other extensions. For brevity, we omit the details and only offer brief discussions. First, the model can be extended to capture “veto” votes, a property we observed in the survey. Specifically, suppose that the value of the project is given by $V = \theta_F(\theta_1 + \dots + \theta_M)$, where θ_F is distributed Bernoulli, and is supposed to capture fraudulent behavior. The idea is that no matter how good the characteristics of a company, if it is fraudulent ($\theta_F = 0$) then the investment is worthless. In addition to observing θ_i , each partner gets a binary signal, denoted s_i , about the project being fraudulent. If the project is not fraudulent ($\theta_F = 1$), then $s_i = 1$ for every i . If the project is fraudulent ($\theta_F = 0$), then $s_i = 0$ with probability z and $s_i = 1$ otherwise, conditionally independent across partners. Intuitively, signal 0 captures an idea of a “red flag” observed by a partner. In this case, allowing the option to veto a project improves decision-making. In equilibrium, a partner vetoes a project if and only if she gets signal 0, a red flag. If a partner observes no red flag, her decision problem is identical to the problem in the basic model (up to a constant multiple). Intuitively, because every voter only considers the cases when her vote is pivotal, she only considers the cases where the veto has not been used. Thus, the results of the baseline model hold (up to a constant multiple), with the added benefit that the model rationalizes the use of veto in voting.

Second, the baseline model can be extended to capture an idea that expertise of a partner is a necessary condition for championing for the project. Specifically, suppose that for any particular investment opportunity only a subset K of the N many partners have expertise in the sector in question and observe characteristics θ_i . The remaining $N - K$ partners are uninformed. In this case, since the prior about the project is that it is non-positive NPV, every uninformed partner does not vote for the project under the

champions voting rule. Then, the analysis of decisions of informed partners is identical to the baseline model with K partners.

Finally, we considered an extension of the baseline model in which each partner i does not observe characteristic θ_i precisely but instead gets a noisy signal S_i about it with errors of signals being independent across partners. This extension is presented in Section C.3 of the appendix. We show that this model is equivalent to the baseline model with the difference that the distribution of signals is a primitive rather than the distribution of characteristics.

4 Empirical Evidence and Quantitative Importance

4.1 Empirical Evidence

Our model suggests that the champions rule dominates other voting rules, but only in settings where the distributions of fundamental characteristics has heavy right tails. The model can therefore also rationalize the survey evidence that VCs start with a champions rule for early-stage investments and migrate to more conventional voting rules for later-stage investments. This would be the model's prediction if the distribution of characteristics for early-stage had heavy right tails, but that this was not as true for late stage investments.

While it is not possible to validate the distribution of individual characteristics of a startup, we are able to examine the *ex post returns* across a wide cross section of venture capital investments. If we find that the returns of later stage investments are less skewed, this would certainly be consistent with the premise the fundamental characteristics for these investments are potentially also less skewed. In the model, if characteristics are subexponentially distributed, then the return distribution is also subexponential with the

same tail as the characteristic distribution. Thus, the empirical evidence on the return distribution across different stages will be consistent with the premise of the model.

Systematic data on returns at the investment level is not available from standard datasets. We received anonymized data on round level returns from Correlation Ventures, a venture capital firm that collects and makes investments in venture capital startups based on quantitative investment strategies. As such, they have a strong incentive to collect, improve and validate the quality of the data they get from standard commercial databases.

The data filter used for the analysis was to first select startups whose headquarters were in the US and had received at least one round of institutional venture capital financing between January 1 2006 and December 31, 2015, and had a realized exit by December 31, 2019. We were provided data on 19,882 rounds of financing in this period with labels corresponding to whether the round of financing was “Seed”, “Series A”, “Series B”, “Series C” or “Series D+”. In other words, while there were slightly more than 19,882 rounds of financing, the rounds including Series D and beyond were aggregated together. Correlation Ventures imputes multiples where these are missing, but for the analysis we conduct, we focus on the subset of 8,603 rounds of financing where the multiple is not missing.

For these rounds, we show the distribution of multiples in Figure 6. As can be seen, nearly 50% of all (non-imputed) returns at the round-level are zero, consistent with the very high failure rates reported in [Kerr, Nanda, and Rhodes-Kropf \(2014\)](#). Table 1 breaks the returns by round, further aggregating those that are in Series C and beyond into a “Series C+” bucket. As can be seen from this, the Seed rounds (and to some extent Series A) stands apart from the other rounds in terms of the skewness of returns. While the median round in the later stages returns a gross return multiple of 0.3, the gross

return for even the 75th percentile Seed round is zero. On the other hand, looking at the 99th percentile shows that which the 99th percentile Series C+ investment returns a gross return of 20, the 99th percentile gross return of a Seed round of financing in the data is 125.

In Figure 7 we compare the distributions of return multiples for each round of financing. For a given point (percentile) in the distribution, we divide the return multiple for each round at that percentile by the return multiple for the overall dataset at the same percentile. We then plot these ratios for all points in the distribution. If the resulting curve for a round has a positive slope, it is because that round has a fatter right tail than the overall distribution. If the resulting curve has a negative slope, it is because that round has a thinner right tail than the overall distribution. The level of the curves (as opposed to the slope) indicate the average return multiple of the round as compared to the overall distribution, and center around 1. We find that Seed is the steepest upward sloping, followed by Series A, while the slopes of Series B and Series C+ are decreasing - indicating that the distribution of returns has the fattest tail at the earliest stage and is less so as the rounds progress.

We also empirically test the return data to estimate the tail index of their respective distributions. The first estimate we use is the truncated Hill's estimate.¹⁵ It is defined as

$$\hat{\alpha}_{Hill} \equiv \frac{n}{\sum_{k=1}^n \log(X_k) - \log(X_n)}$$

where n is the number of observations in the truncated distribution (we truncate the distribution of each return at the highest 5%, 10%, and 20% of returns), and X_k is the return of observation k . These coefficients are plotted in Table 2. The estimated

¹⁵For implementation guidance and advice for these tests, we are grateful to Rustam Ibragimov and Ibragimov, Ibragimov, and Walden (2015).

tail indices for the Seed round are statistically significantly lower than any of the other rounds for all three cutoffs chosen, indicating that the seed round return distribution is the heaviest-tailed of the four. The other three distributions are ranked by stage (later stages have higher coefficients) though the results are not always statistically significant. We also use the method (and Matlab code) of [Huisman et al. \(2001\)](#) to account for the potential of small sample issues in these estimates, and we attain qualitatively similar outcomes as shown in Table 3.

The second estimate we use is the coefficient from a log-log rank-size regression. Here we order the observed returns by size, truncate each distribution to restrict to only positive returns and sort observations by returns. We then estimate the following regression:

$$\log(k) \approx A - \alpha \log(X_k)$$

where k is the rank of a return X_k . The results here are reported in Table 4. Again the estimated α s increase in the stage of the investment, where the seed stage has the lowest estimated α by a statistically significant margin, while the $C+$ round has the highest estimated α .

Our results on the round level returns therefore provide suggestive evidence that is consistent with early stage – and in particular seed stage – fundamental characteristics being much more skewed than those at later (Series C and beyond) stages. Thus, the model could potentially rationalize the use of champions rule for seed and early-stage investments as well as a shift towards a majority voting rule for later stage investments.

4.2 Quantitative Example

To assess quantitative importance of alternative decision making rules, we analyze a numerical example of the model, whose inputs fit the VC context reasonably well. Specifically, we assume that individual characteristics are driven from Pareto (power law) distribution with tail parameter 1.7, which is close to our estimate of the Hill’s tail index for seed investments of 1.62. This implies that the right tail of the return distribution also follows Pareto (power law) distribution with the same tail parameter 1.7. The committee consists of five members. The total number of relevant characteristics is 20, which implies that at most the committee can learn a quarter of the value-relevant information about the project.¹⁶ The other parameter of the Pareto distribution is calibrated so that under the champions rule 1% of projects are accepted on average. Regardless of its quality, each project is assumed to fail with probability 50% returning zero payoff and to return a non-zero payoff equal to the sum of characteristics with probability 50%. Overall, this distribution means that a typical project considered by the committee is clearly bad, but a very small fraction of projects are exceptionally promising. We keep the assumption that characteristics observed by partners are independent.¹⁷ Three decision rules are compared in this setting: (1) champions (one positive vote out of five is needed for investment); (2) majority (three out five); (3) unanimity (five out of five).

This example produces the following results. In this example, the champions rule is optimal, yielding higher values to the partnership than the alternatives. More interestingly, it leads to investment in projects with a very different profile than the majority and unanimity rules. Projects funded under the champions rule have significantly higher av-

¹⁶We have also analyzed quantitative examples with five and ten relevant characteristics and five partners. In these cases, the probability that the champions rule catches an outlier is higher than with 20 characteristics, but the conceptual comparison with the other decision rules is unchanged.

¹⁷The assumption of independent characteristics exaggerates the difference between the champions rule.

erage payoffs, higher variance, and higher skewness. For example, the standard deviation of payoffs for the champions, majority, and unanimity rules are 4.50, 1.42, and 0.96 per dollar of investment, respectively. Skewness (also normalized per dollar of investment) are 60.44, 3.65, and 0.29, respectively. The 95th quantile of the realized per-investment multiple (including both zero and positive payoffs) is 4.78 for the champions rule, 3.10 for majority, and 2.24 for unanimity. The corresponding numbers for the 99th quantile are 10.61, 5.54, and 3.15, respectively.¹⁸ In other words, the payoff distribution is right-skewed under any decision rule, which is simply due to the nature of investments, but it is significantly more right-skewed under the champions rule. Different project profile implies that the probability of “catching a unicorn” is significantly higher for the champions rule than the alternatives. The model-implied probability of getting a realized payoff per investment exceeding multiple 10x, adjusted for the time value of money, is 1.12% for the champions rule and 0.21% for the majority rule. The model-implied probability of getting a realized payoff per investment exceeding multiple 10x, unadjusted for the time value of money,¹⁹ is 2.89% for the champions rule and 0.7% for the majority rule. For example, if we consider a fund with 25 uncorrelated investments, then the probability that it least one of these investments will deliver such a multiple is 52% for the champions rule and 16.1% for the majority rule.

¹⁸Compared to the example in Section 3, in this example, each investment project returns zero with probability 50%. Thus, the equilibrium voting thresholds and return distributions are different from the numbers in Table 5.

¹⁹Assuming a 5-year investment horizon and the discount rate of 12%, this implies the adjusted multiple of 6.2.

5 Conclusion

We provide novel empirical evidence on the voting practices of venture capital investors in the US, showing that investors use different voting rules for different types of investments, and importantly different voting rules than would appear to be optimal based on received wisdom. For early-stage investments, they tend to favor champion voting rules, where one partner can unilaterally make the decision to invest, while for later stage investments, their strategies shift towards more consensus based rules, like majority or unanimity. When asked why they do it, venture capital investors overwhelmingly say that the champion voting rule maximizes the chances of ‘catching outliers’. We next analyze this survey evidence through the prism of existing canonical models of information aggregation in committees. We conclude that the model that is most consistent with the survey evidence is one in which the value of the project depends on multiple project attributes, an extremely strong attribute can overcompensate for the presence of other weak attributes, and each partner gets a signal about one attribute but not the others. In this model, distributions with heavy right tails (similar to the distribution of early-stage investments) are ones where a champions voting rule is optimal. In contrast, later stage investments where distributions have significantly less fat tails imply that ‘majority’ or ‘unanimous’ voting rules are optimal. We then confirm that these conditions, which justify champions voting rules, are satisfied by the ex-post returns of early-stage investments, while they are not by later stage investments, thus validating the voting rules seen in practice.

While venture capital provides a useful empirical setting in which to test and validate the model, the implications of the paper are more wide-ranging: effectively any decision made by a group could benefit from lower levels of consensus if the options being considered satisfy a few conditions: first, that they are not ex-ante perfectly observable; second

that the members of the group cannot perfectly communicate their information to one another; third that the ex-ante distribution of outcomes for each option has a sufficiently large right tail. This type of setting could be applicable to hiring decisions, research and development, innovation, financial investment, and more.

References

- Azevedo, Eduardo M., Alex Deng, Jose Luis Montiel Olea, Justin Rao, and E. Glen Weyl. 2020. “A/B testing with fat tails.” *Journal of Political Economy* 128:4614–4672.
- Bergemann, Dirk and Ulrich Hege. 1998. “Venture capital financing, moral hazard, and learning.” *Journal of Banking & Finance* 22 (6-8):703–735.
- Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. “Are Ideas Getting Harder to Find?” *American Economic Review* 110:1104–1144.
- Bond, Philip and Hülya Eraslan. 2010. “Strategic voting over strategic proposals.” *The Review of Economic Studies* 77 (2):459–490.
- Chan, Jimmy, Alessandro Lizzeri, Wing Suen, and Leeat Yariv. 2018. “Deliberating collective decisions.” *The Review of Economic Studies* 85 (2):929–963.
- Condorcet, Marquis de. 1785. “Essay on the Application of Analysis to the Probability of Majority Decisions.” *Paris: Imprimerie Royale* .
- Cornelli, Francesca and Oved Yosha. 2003. “Stage financing and the role of convertible securities.” *Review of Economic Studies* 70 (2):1–32.
- Coughlan, Peter J. 2000. “In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting.” *American Political science review* :375–393.
- Csaszar, Felipe A. and J. P. Eggers. 2013. “Organizational Decision Making: An Information Aggregation View.” *Management Science* 59 (10):2257–2277.
- DeMarzo, Peter M., Ilan Kremer, and Andrzej Skrzypacz. 2005. “Bidding with securities: auctions and security design.” *American Economic Review* 95:936–959.

- Eisenmann, Thomas and Liz Kind. 2014. “Andreessen Horowitz.” *Harvard Business School Case* 814-060.
- Feddersen, Timothy and Wolfgang Pesendorfer. 1997. “Voting behavior and information aggregation in elections with private information.” *Econometrica: Journal of the Econometric Society* :1029–1058.
- . 1998. “Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting.” *American Political science review* :23–35.
- Gerling, Kerstin, Hans Peter Gruner, Alexandra Kiel, and Elisabeth Schulte. 2005. “Information acquisition and decision making in committees: A survey.” *European Journal of Political Economy* 21:563–597.
- Goldstein, Itay and Liyan Yang. 2017. “Information disclosure in financial markets.” *Annual Review of Financial Economics* 9:101–125.
- Gompers, Paul. 1995. “Optimal Investment, monitoring, and the staging of venture capital.” *Journal of Finance* 50:1461–1490.
- Gompers, Paul, William Gornall, Steven Kaplan, and Ilya Strebulaev. 2020. “How do venture capitalists make decisions?” *Journal of Financial Economics* 135 (1):169–190.
- Graham, John and Campbell Harvey. 2001. “The theory and practice of corporate finance: evidence from the field.” *Journal of Financial Economics* 60:187–243.
- Hall, Robert E and Susan E Woodward. 2010. “The Burden of the Nondiversifiable Risk of Entrepreneurship.” *American Economic Review* 100 (3):1163–1194.
- Harris, Milton and Artur Raviv. 2005. “Allocation of decision-making authority.” *Review of Finance* 9:353–383.

- . 2008. “A theory of board control and size.” *Review of Financial Studies* 21:1797–1832.
- Hellmann, Thomas. 1998. “The allocation of control rights in venture capital contracts.” *RAND Journal of Economics* 29:57–76.
- Hellmann, Thomas and Manju Puri. 2000. “The interaction between product market and financing strategy: The role of venture capital.” *Review of Financial Studies* 13:959–984.
- . 2002. “Venture capital and the professionalization of start-up firms: Empirical evidence.” *Journal of Finance* 57:169–197.
- Hellwig, Martin F. 1980. “On the aggregation of information in competitive markets.” *Journal of Economic Theory* 22:477–498.
- Huisman, Ronald, Kees G Koedijk, Clemens J M Kool, and Franz Palm. 2001. “Tail-index estimates in small samples.” *Journal of Business & Economic Statistics* 19 (2):208–216.
- Ibragimov, Marat, Rustam Ibragimov, and Johan Walden. 2015. *Heavy-tailed distributions and robustness in economics and finance*, vol. 214. Springer.
- Jackson, Matthew O and Xu Tan. 2013. “Deliberation, disclosure of information, and voting.” *Journal of Economic Theory* 148 (1):2–30.
- Kaplan, Steven N, Berk Sensoy, and Per Stromberg. 2009. “Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies.” *Journal of Finance* 61 (4):75–115.

- Kaplan, Steven N and Per Strömberg. 2003. “Financial contracting theory meets the real world: An empirical analysis of venture capital contracts.” *The review of economic studies* 70 (2):281–315.
- Kerr, William R, Ramana Nanda, and Matthew Rhodes-Kropf. 2014. “Entrepreneurship as experimentation.” *The Journal of Economic Perspectives* 28 (3):25–48.
- Khanna, Naveen and Richmond D. Mathews. 2011. “Can herding improve investment decisions?” *RAND Journal of Economics* 42 (1):150–174.
- Kortum, Samuel and Josh Lerner. 2000. “Assessing the Contribution of Venture Capital to Innovation.” *The RAND Journal of Economics* 31 (4):674–692.
- Kremer, Michael. 1993. “The O-Ring Theory of Economic Development.” *Quarterly Journal of Economics* 108 (3):551–575.
- Ladha, Krishna K. 1992. “The Condorcet jury theorem, free speech, and correlated votes.” *American Journal of Political Science* :617–634.
- Lerner, Josh. 1995. “Venture Capital’s Role in Financing Innovation: What We Know and How Much We Still Need to Learn.” *Journal of Finance* 50:301–318.
- Lerner, Josh and Ramana Nanda. 2020. “Venture Capital’s Role in Financing Innovation: What We Know and How Much We Still Need to Learn.” *Journal of Economic Perspectives* 34 (3).
- Malenko, Nadya. 2014. “Communication and Decision-Making in Corporate Boards.” *Review of Financial Studies* 27 (5):1486–1532.
- Manso, Gustavo. 2011. “Motivating Innovation.” *Journal of Finance* 66 (5):1823–1860.

- McLennan, Andrew. 1998. "Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents." *American Political Science Review* 92:413–418.
- Moldovanu, Benny and Xianwen Shi. 2013. "Specialization and Partisanship in Committee Search." *Theoretical Economics* 8:751–774.
- Name-Correa, Alvaro J and Huseyin Yildirim. 2019. "Social pressure, transparency, and voting in committees." *Journal of Economic Theory* 184:104943.
- Sah, Raaj Kumar and Joseph E. Stiglitz. 1986. "The Architecture of Economic Systems: Hierarchies and Polyarchies." *American Economic Review* 76:716–727.
- . 1988. "Committees, Hierarchies and Polyarchies." *Economic Journal* 98:451–470.
- Scharfstein, David and Jeremy Stein. 2000. "The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment." *The Journal of Finance* 55 (5):2537–2564.
- Scherer, F. M. and Dietmar Harhoff. 2000. "Technology policy for a world of skewed distributed outcomes." *Research Policy* 29 (4-5):559–566.
- Thiel, Peter and Blake Masters. 2014. *Zero to One*. Crown Business.
- Verrecchia, Robert E. 1982. "Information acquisition in a noisy rational expectations economy." *Econometrica* 50:1415–1430.
- Yung, Chris. 2005. "IPOs with Buy- and Sell-Side Information Production: The Dark Side of Open Sales." *Review of Financial Studies* 18 (1):327–347.

A Figures and Tables

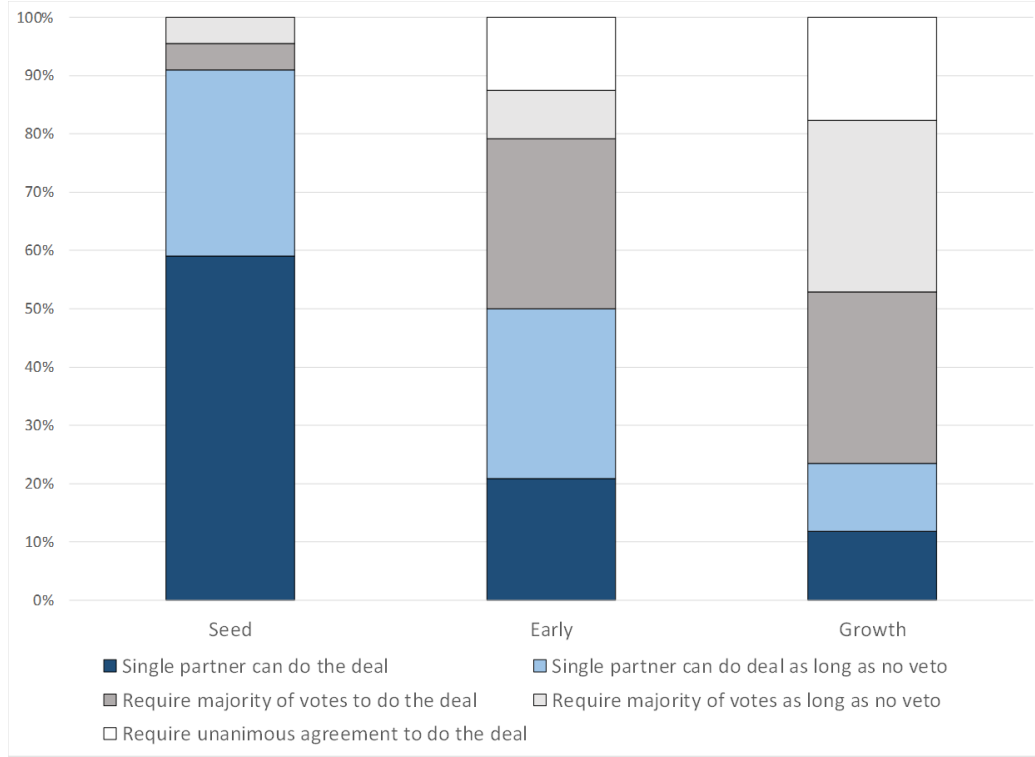


Figure 1: Breakdown of Formal Voting by stage of investment: This figure shows the results from survey question 3 from Figure 12. The three bars correspond, for left to right, to the voting rules for seed, early and growth investment rounds. The bars all add up to 100% and are broken down into different voting rules.

	Percentile					
	25th	50th	75th	90th	95th	99th
Seed	0.00	0.00	0.00	3.25	8.81	124.58
Series A	0.00	0.00	2.56	8.80	20.15	77.65
Series B	0.00	0.26	2.56	6.56	11.14	36.10
Series C+	0.00	0.30	1.87	4.01	6.23	19.76
Total Sample	0.00	0.00	2.13	5.98	11.21	49.71

Table 1: Return multiple percentiles for overall data, as well as by round of financing.

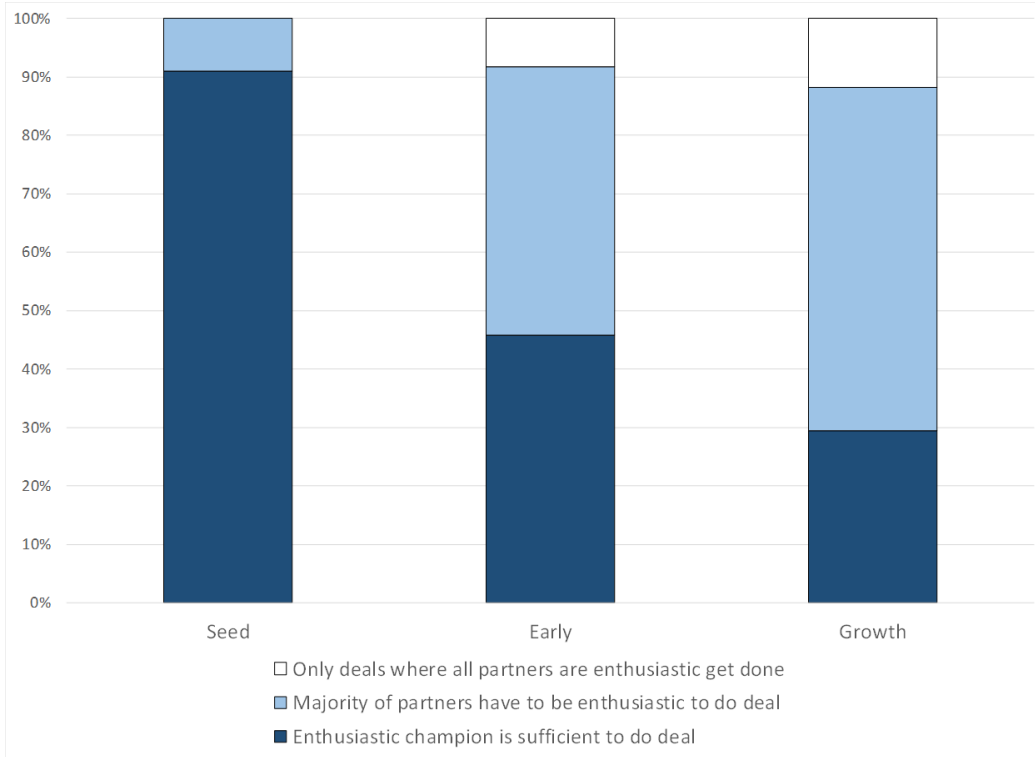


Figure 2: Breakdown of Informal Voting by stage of investment: This figure shows the results from survey question 4 from Figure 12. The three bars correspond, for left to right, to the voting rules for seed, early and growth investment rounds. The bars all add up to 100% and are broken down into different voting rules.

	Hill's Tail Index Estimate			
	Seed Round	Series A	Series B	Series C+
20% Cutoff Estimate	0.62 (0.055)	2.00 (0.09)	2.64 (0.11)	3.12 (0.13)
10% Cutoff Estimate	1.53 (0.19)	2.24 (0.15)	3.01 (0.18)	3.26 (0.20)
5% Cutoff Estimate	1.62 (0.29)	2.76 (0.26)	3.03 (0.25)	3.04 (0.18)

Table 2: This table shows the Hill estimator of the tail index for each round, along with standard errors for each round.

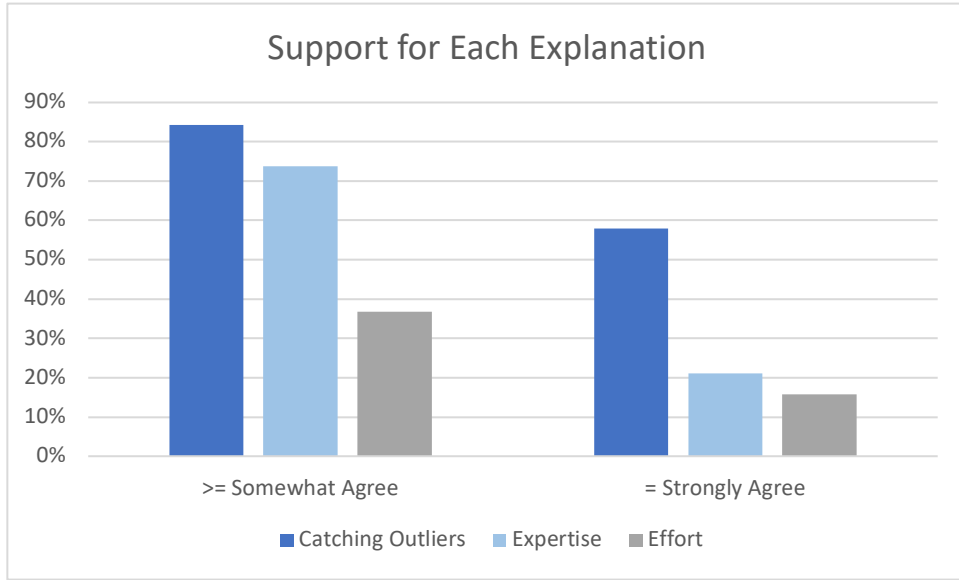


Figure 3: This figure shows the support for each suggested explanation for champions rule voting. The left hand bars correspond to the fraction of respondents who gave each explanation their highest (possibly joint highest) rating. The right hand bars correspond to the fraction of respondents who gave each explanation their strongest, exclusive support.

	Small Sample Tail-Index Estimate			
	Seed Round	Series A	Series B	Series C+
Small Sample Estimate	0.1978 (0.0128)	0.7718 (0.0007)	0.6901 (0.0006)	0.7380 (0.0006)

Table 3: This table shows the Tail-index estimate in small samples of [Huisman et al. \(2001\)](#), using a 20.7% cutoff.

	Log-Log Tail Index Estimate			
	Seed Round	Series A	Series B	Series C+
Coefficient Estimate	0.35 (0.02)	0.49 (0.009)	0.49 (0.009)	0.61 (0.01)

Table 4: This table shows the coefficients from the log-log rank-size regressions for each round, along with standard errors for each round.

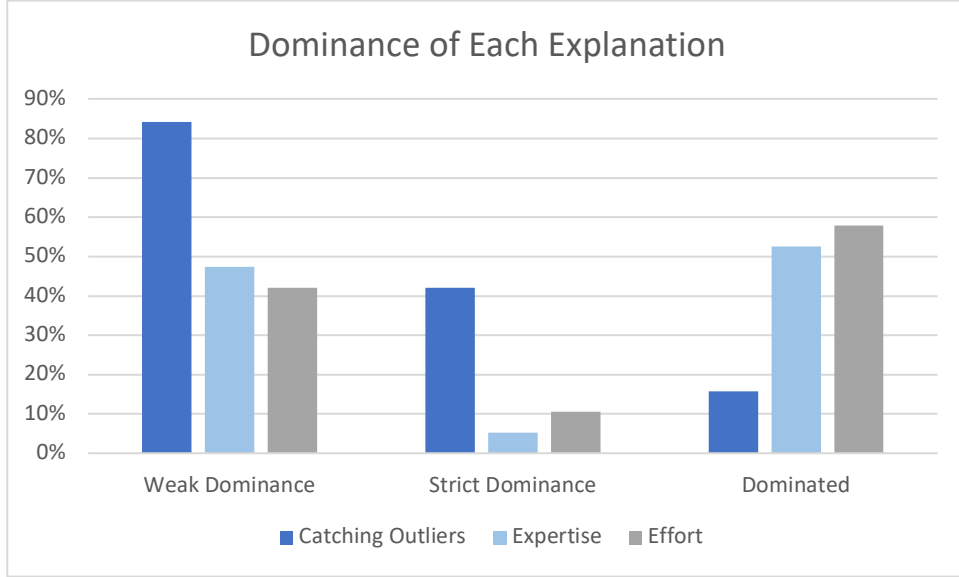


Figure 4: This figure shows the relative dominance for each suggested explanation for champions rule voting. The first set of bars correspond to the fraction of respondents that rated no other explanation more highly; the second set of bars correspond to the fraction of respondents that rated no other explanation as highly; the last set of bars correspond to the fraction of respondents that rated other explanations more highly.

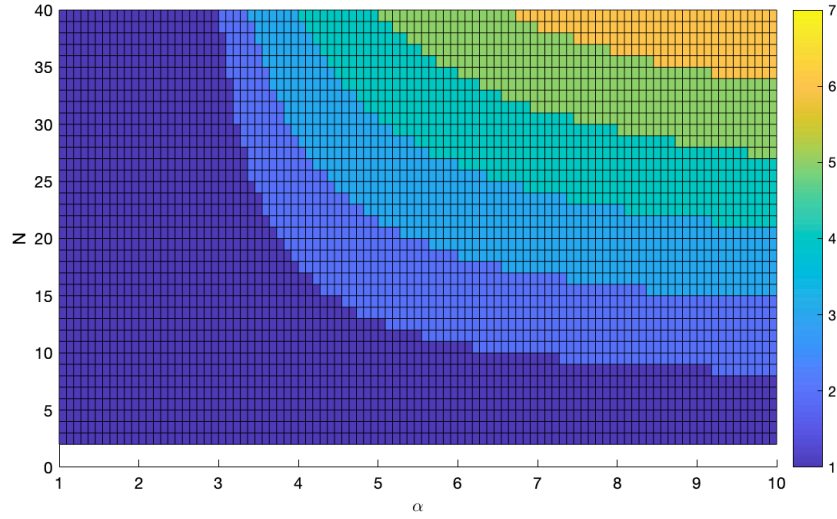


Figure 5: This figure shows the optimal voting rule k as a function of the number of investors/characteristics N , under the assumption that the characteristics are distributed according to a Pareto distribution with shape parameter α . Darker colors correspond to lower values of k (the dark blue corresponds to champions rule, while lighter colors correspond to higher values of k).

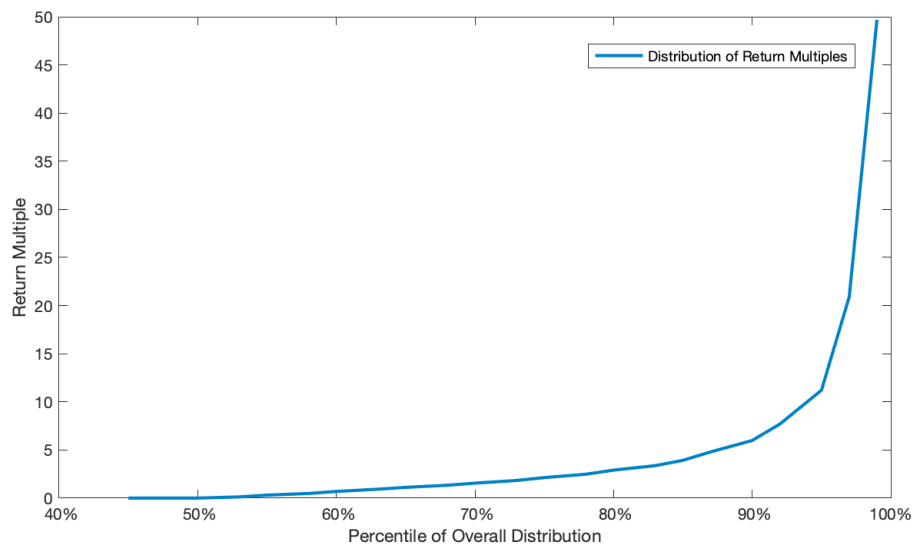


Figure 6: This figure shows the distribution of return multiples for the total sample of investments.

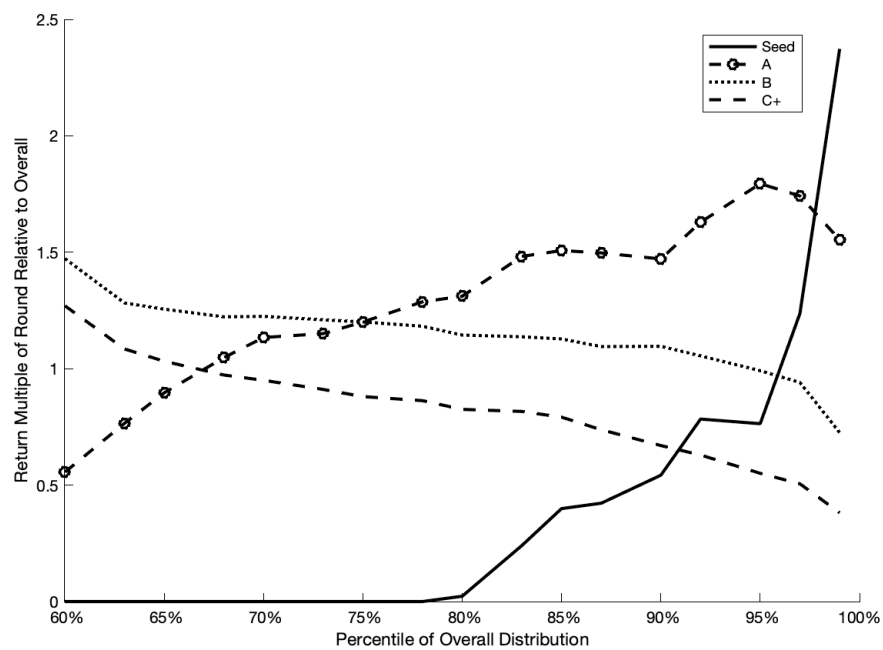


Figure 7: This figure shows the ratio of return multiples for a given round to the total sample of investments.

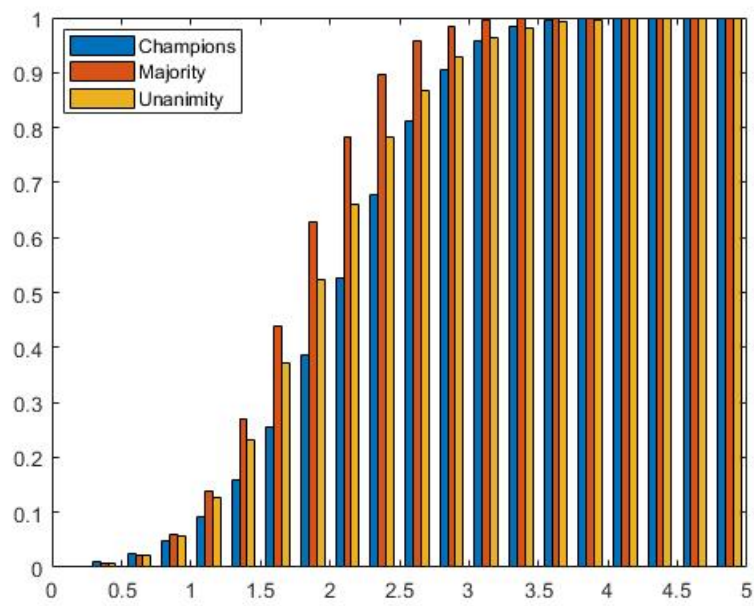


Figure 8: This figure shows the distribution of the probability of accepting different projects under three different voting rules, when the information setup is truth plus noise.

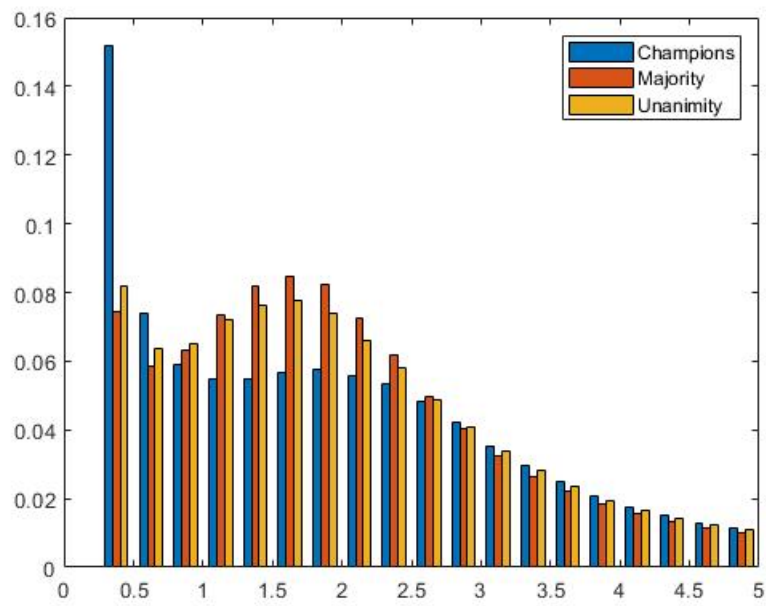


Figure 9: This figure shows the distribution of ex-post returns of projects under three different voting rules, when the information setup is truth plus noise.

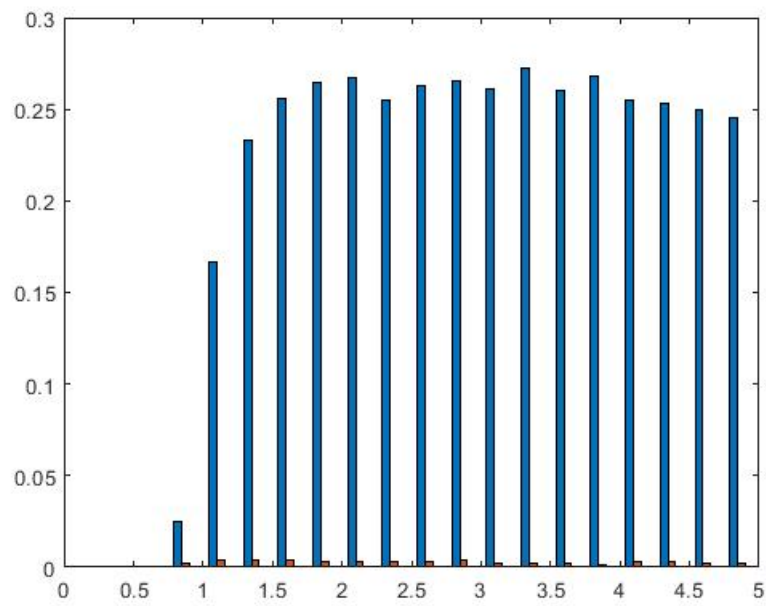


Figure 10: This figure shows the distribution of the probability of accepting different projects under three different voting rules, when the information setup is the sum of signals.

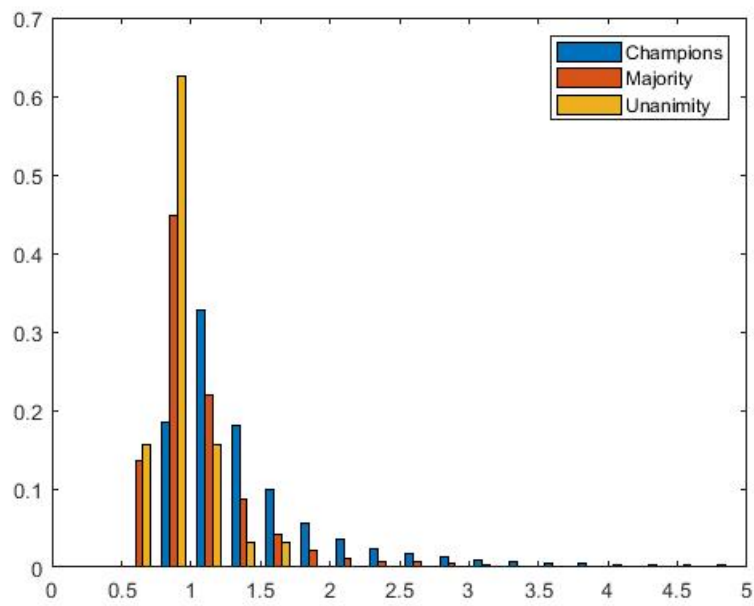


Figure 11: This figure shows the distribution of ex-post returns of projects under three different voting rules, when the information setup is the sum of signals.

Truth + Noise Setup			
percentile	Champions	Majority	Unanimity
25th	85%	119%	114%
50th	196%	194%	196%
75th	330%	305%	319%
90th	568%	525%	550%
95th	855%	790%	827%
99th	2209%	2050%	2142%
Additive Signals Setup			
percentile	Champions	Majority	Unanimity
25th	104%	81%	78%
50th	124%	95%	89%
75th	163%	117%	95%
90th	239%	155%	112%
95th	333%	195%	127%
99th	779%	365%	162%

Table 5: This table shows the quantiles of the distribution of gross returns for different voting rules under two different informational setups.

B Appendix

B.1 Proofs

Proof of Proposition 1 Consider the distribution of project values V , conditional on $V \geq \hat{V}$ for a very large \hat{V} . By the catastrophe principle of subexponential distributions, this distribution is similar to the distribution of $\max(\theta_1, \dots, \theta_N)$:

$$\Pr(\theta_1 + \theta_2 + \dots + \theta_M > \hat{V}) \sim \Pr(\max(\theta_1, \dots, \theta_M) > \hat{V}).$$

Intuitively, extreme project values are driven by extreme realizations of one characteristic rather than by very high realizations of many characteristics. There are two possible cases. The first case is $\max(\theta_1, \dots, \theta_N) = \max(\theta_1, \dots, \theta_M)$, i.e., the extreme characteristic is learned by one of the N partners. The second case is $\max(\theta_1, \dots, \theta_N) < \max(\theta_1, \dots, \theta_M)$, i.e., the extreme characteristic is not learned by one of the N partners. By independence, these cases occur with probabilities $\frac{N}{M}$ and $\frac{M-N}{M}$, respectively.

Consider the first case. Given a voting rule k , the probability that a project with extreme value gets rejected is

$$\begin{aligned} & \lim_{\hat{V} \rightarrow \infty} \Pr(d_k(\theta) = 0 | V > \hat{V}, \max(\theta_1, \dots, \theta_N) = \max(\theta_1, \dots, \theta_M)) \\ &= \lim_{\hat{V} \rightarrow \infty} \Pr(d_k(\theta) = 0 | \max(\theta_1, \dots, \theta_N) > \hat{V}) \\ &= \lim_{\hat{V} \rightarrow \infty} \Pr\left(\sum_{i=2}^N v_i < k-1 | \theta_i < \theta_1, i \in \{2, \dots, N\}, \theta_1 > \hat{V}\right) \\ &= \begin{cases} 0, & \text{if } k = 1, \\ \sum_{m=0}^{k-2} C_{N-1}^m (1 - F(\hat{\theta}_k))^m F(\hat{\theta}_k)^{N-1-m}, & \text{if } k \geq 2. \end{cases} \end{aligned}$$

The first equality holds by the catastrophe principle. The second equality holds by the fact that the project gets rejected if and only if out of $N-1$ partners who do not get an extreme signal,

there are fewer than $k - 1$ votes “for” (the partner that gets an extreme signal always votes “for”). Finally, the last equality holds by the fact that the distribution of each θ_i , $i \in \{2, \dots, N\}$, conditional on $\max\{\theta_1, \dots, \theta_N\} = \theta_1$, is the distribution of θ_i , truncated at θ_1 , and as θ_1 becomes infinite, it approaches the distribution of θ_i . Thus, the probability of rejection is zero when $k = 1$ and strictly positive when $k > 1$. Therefore, conditional on the maximum characteristic being learned about (which occurs with probability $\frac{N}{M}$), the champions rule never rejects high value projects, while any voting rule $k > 1$ rejects them with positive probability.

Consider the second case in which $\max(\theta_1, \dots, \theta_N) < \max(\theta_1, \dots, \theta_M)$. This case occurs with probability $\frac{M-N}{M}$. In this case,

$$\begin{aligned} & \lim_{\hat{V} \rightarrow \infty} \Pr(d_k(\theta) = 0 | V > \hat{V}, \max(\theta_1, \dots, \theta_N) < \max(\theta_1, \dots, \theta_M)) \\ &= \lim_{\hat{V} \rightarrow \infty} \Pr(d_k(\theta) = 0 | \max(\theta_{N+1}, \dots, \theta_M) > \hat{V}) \\ &= \Pr(d_k(\theta) = 0). \end{aligned}$$

The first equality holds by the catastrophe principle, and the second equality holds by independence of all M characteristics.

Adding up the two cases, we obtain that the probability of catching a “superstar” for voting rule k is

$$\frac{N}{M} + \frac{M-N}{M} \Pr(d_1(\theta) = 1) \quad (4)$$

for the champions rule ($k = 1$), and

$$\frac{N}{M} \left(\sum_{m=k-1}^{N-1} C_{N-1}^m (1 - F(\hat{\theta}_k))^m F(\hat{\theta}_k)^{N-1-m} \right) + \frac{M-N}{M} \Pr(d_k(\theta) = 0) \quad (5)$$

for any other voting rule k .

Clearly, if $\Pr(d_k(\theta) = 1) \leq \Pr(d_1(\theta) = 1)$, then (4) exceeds (5), as the first term is strictly higher and the second term is weakly higher. Alternatively, if $\Pr(d_k(\theta) = 1) > \Pr(d_1(\theta) = 1)$,

then (4) exceeds (5) if

$$\frac{N}{M-N} \left(\sum_{m=0}^{k-2} C_{N-1}^m \left(1 - F(\hat{\theta}_k) \right)^m F(\hat{\theta}_k)^{N-1-m} \right) > \Pr(d_k(\theta) = 1) - \Pr(d_1(\theta) = 1). \quad (6)$$

In particular, this condition is always satisfied if $\Pr(d_1(\theta) = 1) \geq \Pr(d_k(\theta) = 1)$. Alternatively, for any $\Pr(d_k(\theta) = 1) - \Pr(d_1(\theta) = 1)$, (6) is satisfied if the ratio of learned to unknown characteristics is sufficiently high.

Proof of Proposition 2. We start by showing that the model with $M > N$ is equivalent to the model with $M = N$ and a modified investment cost. Then, given this result, it is sufficient to prove this and the subsequent propositions for the case of $M = N$ only.

Let us decompose the value of the project into a potentially learned part and an unknown part:

$$V = V_1 + V_2 - I,$$

where $V_1 = \sum_{i=1}^N \theta_i$ and $V_2 = \sum_{i=N+1}^M \theta_i$. By independence of θ_i , the expected value of any decision rule k is equal to the expected value of the same decision rule if each project's value is given by

$$\tilde{V} = V_1 - \tilde{I},$$

where $\tilde{I} = I - (M - N) E[\theta]$ is the investment cost modified by the expected value of the unlearned characteristics. Indeed:

$$\begin{aligned} & \mathbb{E} \left[\left(\sum_{i=1}^M \theta_i - I \right) d_k(\theta_1, \dots, \theta_N) \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\left(\sum_{i=1}^M \theta_i - I \right) d_k(\theta_1, \dots, \theta_N) \mid \theta_1, \dots, \theta_N \right] \right] \\ &= \mathbb{E} \left[\left(\sum_{i=1}^N \theta_i - \tilde{I} \right) d_k(\theta_1, \dots, \theta_N) \right], \end{aligned}$$

where we use $d_k(\theta_1, \dots, \theta_N) \in \{0, 1\}$ to denote the investment decision under voting rule k when

the realized vector of learned characteristics is $(\theta_1, \dots, \theta_N)$.

Given this result, without loss of generality, consider the case $M = N$. Let $U^*(k)$ be the expected value from voting rule k . We can write it as:

$$U^*(k) = \sum_{m=k}^N C_N^m \left(1 - F(\hat{\theta}_k)\right)^m F(\hat{\theta}_k)^{N-m} \left(m \mathbb{E}[\theta | \theta \geq \hat{\theta}_k] + (N-m) \mathbb{E}[\theta | \theta < \hat{\theta}_k] - I\right) \quad (7)$$

Note that

$$\mathbb{E}[\theta] = F(\hat{\theta}_k) \mathbb{E}[\theta | \theta < \hat{\theta}_k] + (1 - F(\hat{\theta}_k)) \mathbb{E}[\theta | \theta \geq \hat{\theta}_k].$$

If $\int_{\hat{\theta}_1}^{\infty} \theta dF(\theta)$ is arbitrarily close to $\mathbb{E}[\theta]$, then $\int_0^{\hat{\theta}_1} \theta dF(\theta)$ is arbitrarily close to zero. Consider (7) for $k = 1$. Since $1 - F(\hat{\theta}_1) \rightarrow 0$ and $(1 - F(\hat{\theta}_1)) \mathbb{E}[\theta | \theta \geq \hat{\theta}_1] \rightarrow \mathbb{E}[\theta]$, all terms in the summand (7) other than $m = 1$ are infinitesimal. Hence,

$$U^*(1) \rightarrow N \left(1 - F(\hat{\theta}_1)\right) \left(\mathbb{E}[\theta | \theta \geq \hat{\theta}_1] - I\right) = N \mathbb{E}[\theta].$$

This limit is the maximal expected value that the decision-maker with perfect knowledge of the project value could realize. In contrast, for any $k > 1$, $U^*(k)$ is bounded away from $N \mathbb{E}[\theta]$, since a project with $\theta_i > \hat{\theta}_1$ for one of $i \in \{1, \dots, N\}$ is missed with strictly positive probability. Hence, $k = 1$ is optimal.

Proof of Proposition 3. By the result at the beginning of the proof of the previous proposition, it is sufficient to consider the case of $M = N$.

Let $U(k, \hat{\theta})$ be the expected value from a decision rule that requires k approval votes for investment, if each if each partner votes for if and only if her signal exceeds $\hat{\theta}$. We can re-write

it as:

$$\begin{aligned}
U(k, \hat{\theta}) &= \sum_{m=k}^N C_N^m \left(1 - F(\hat{\theta})\right)^m F(\hat{\theta})^{N-m} \left(m \mathbb{E}[\theta | \theta \geq \hat{\theta}] + (N-m) \mathbb{E}[\theta | \theta < \hat{\theta}] - I\right) \\
&= \sum_{m=k}^N \left(1 - F(\hat{\theta})\right)^m F(\hat{\theta})^{N-m} \\
&\quad \times \left(\frac{N!}{(m-1)!(N-m)!} \frac{\int_{\hat{\theta}}^{\infty} \left(\theta - \frac{I}{N}\right) dF(\theta)}{1 - F(\hat{\theta})} + \frac{N!}{m!(N-m-1)!} \frac{\int_0^{\hat{\theta}} \left(\theta - \frac{I}{N}\right) dF(\theta)}{F(\hat{\theta})} \right) \\
&= (N \mathbb{E}[\theta] - I) \left(1 - I_{F(\hat{\theta})}(N-k, k)\right) + C_{N-1}^{k-1} \left(1 - F(\hat{\theta})\right)^{k-1} F(\hat{\theta})^{N-k} \left(\int_{\hat{\theta}}^{\infty} (N\theta - I) dF(\theta)\right),
\end{aligned}$$

where $C_N^m = N! / (m!(N-m)!)$ and $I_x(a, b)$ is the regularized incomplete beta function. Let $U^*(k)$ be the equilibrium value of the project under decision rule k , i.e., $U(k, \hat{\theta})$ evaluated at the equilibrium voting cut-off $\hat{\theta}$. Since all committee members have the same utility functions, this is the game of common interest, and thus $\hat{\theta}$ that maximizes $U(k, \hat{\theta})$ constitutes an equilibrium (McLennan (1998)). Thus,

$$U^*(k) = \max_{\hat{\theta}} U(k, \hat{\theta}) = \max_p \bar{U}(k, p), \quad (8)$$

where

$$\begin{aligned}
\bar{U}(k, p) &= (N \mathbb{E}[\theta] - I) (1 - I_{1-p}(N-k, k)) + C_{N-1}^{k-1} p^{k-1} (1-p)^{N-k} \left(\int_{1-p}^1 (Nu(z) - I) dz\right) \\
u(z) &= F^{-1}(z),
\end{aligned}$$

(for convenience, we introduced changes in variables). Intuitively, $p = 1 - F(\hat{\theta})$ is the probability that a committee member votes for the project, implied by threshold $\hat{\theta}$, and $u(z)$ is the value of a characteristic in the z th quantile.

Under Pareto distribution with shape α , $\int_{1-p}^1 (Nu(z) - I) dz = N \frac{\alpha}{\alpha-1} x_m p^{1-\frac{1}{\alpha}} - Ip$, where

$x_m = -l$. Hence, (9) simplifies to

$$\bar{U}(k, p) = (N\mathbb{E}[\theta] - I)(1 - I_{1-p}(N - k, k)) + C_{N-1}^{k-1} p^{k-1} (1 - p)^{N-k} \left(N \frac{\alpha}{\alpha - 1} x_m p^{1-\frac{1}{\alpha}} - Ip \right). \quad (10)$$

We first consider the case in which the project is ex-ante zero-NPV ($N\mathbb{E}[\theta] = I$), and then consider the more realistic case in which it is negative NPV ($N\mathbb{E}[\theta] < I$).

Case 1: $N\mathbb{E}[\theta] = I$. In this case, the first term in (10) equals zero, so (8) simplifies to

$$U^*(k) = C_{N-1}^{k-1} \max_p p^{k-1} (1 - p)^{N-k} \left(N \frac{\alpha}{\alpha - 1} x_m p^{1-\frac{1}{\alpha}} - Ip \right)$$

The zero-NPV condition implies $Nx_m = \frac{\alpha-1}{\alpha} I$. Therefore, $I = N \frac{\alpha}{\alpha-1} x_m$. Hence,

$$U^*(k) = C_{N-1}^{k-1} \max_p p^{k-1} (1 - p)^{N-k} \left(\frac{\alpha}{\alpha - 1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha - 1} p \right) Nx_m.$$

Consider the ratio of $U^*(k+1)$ to $U^*(k)$:

$$\begin{aligned} \frac{U^*(k+1)}{U^*(k)} &= \frac{N-k}{k} \frac{\max_p p^k (1-p)^{N-k-1} \left(\frac{\alpha}{\alpha-1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha-1} p \right)}{\max_p p^{k-1} (1-p)^{N-k} \left(\frac{\alpha}{\alpha-1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha-1} p \right)} \\ &= \frac{N-k}{k} \frac{\max_p p^k (1-p)^{N-k-1} \left(p^{1-\frac{1}{\alpha}} - p \right)}{\max_p p^{k-1} (1-p)^{N-k} \left(p^{1-\frac{1}{\alpha}} - p \right)} \end{aligned}$$

We will show that $\frac{U^*(k+1)}{U^*(k)} < 1$ for all $k \in \{1, \dots, N-1\}$ for $\alpha \rightarrow 1$. Notice that

$$\lim_{\alpha \rightarrow 1} \left(p^{1-\frac{1}{\alpha}} - p \right) = 1 - p \quad \forall p \in (0, 1).$$

Therefore,

$$\begin{aligned}\lim_{\alpha \rightarrow 1} \frac{U^*(k+1)}{U^*(k)} &= \frac{N-k}{k} \frac{\max_p p^k (1-p)^{N-k}}{\max_p p^{k-1} (1-p)^{N-k+1}} = \frac{N-k}{k} \frac{k^k (N-k)^{N-k}}{(k-1)^{k-1} (N-k+1)^{N-k+1}} \\ &= \frac{k^{k-1} (N-k)^{N-k+1}}{(k-1)^{k-1} (N-k+1)^{N-k+1}} = \left(\frac{k}{k-1} \right)^{k-1} \left(\frac{N-k}{N-k+1} \right)^{N-k+1}\end{aligned}$$

Taking a natural logarithm, $\lim_{\alpha \rightarrow 1} \frac{U^*(k+1)}{U^*(k)} < 1$ is equivalent to

$$(k-1)(\ln(k) - \ln(k-1)) + (N-k+1)(\ln(N-k) - \ln(N-k+1)) < 0$$

Consider $f(x) = x \ln\left(1 + \frac{1}{x}\right)$.

$$\begin{aligned}f'(x) &= \ln\left(1 + \frac{1}{x}\right) - \frac{\frac{1}{x}}{1 + \frac{1}{x}} = \ln\left(1 + \frac{1}{x}\right) - \frac{1}{x+1} \\ \lim_{x \rightarrow 0} f'(x) &= \infty \\ f''(x) &= -\frac{x}{x+1} \left(\frac{1}{x^2} \right) + \frac{1}{(x+1)^2} \\ &= \frac{1}{(x+1)^2} - \frac{1}{x(x+1)} = \frac{x - (x+1)}{x(x+1)^2} < 0 \\ \lim_{x \rightarrow \infty} f'(x) &= \lim_{x \rightarrow \infty} \left(\ln\left(1 + \frac{1}{x}\right) - \frac{1}{x+1} \right) = 0\end{aligned}$$

Hence, $f'(x)$ is strictly decreasing in x starting from infinity for $x \rightarrow 0$ to zero for $x \rightarrow \infty$.

Hence, $f'(x) > 0$ for all x . Therefore, $f(x) = x \ln\left(1 + \frac{1}{x}\right)$ is maximized at $x \rightarrow \infty$, in which case

$$\lim_{x \rightarrow \infty} x \ln\left(1 + \frac{1}{x}\right) = 1$$

Consider $g(x) = (x+1) \ln\left(1 + \frac{1}{x}\right)$:

$$\begin{aligned} g'(x) &= \ln\left(1 + \frac{1}{x}\right) - (x+1) \frac{\frac{1}{x^2}}{1 + \frac{1}{x}} \\ &= \ln\left(1 + \frac{1}{x}\right) - \frac{x+1}{x^2+x} = \ln\left(1 + \frac{1}{x}\right) - \frac{1}{x} \\ g''(x) &= -\frac{\frac{1}{x^2}}{1 + \frac{1}{x}} + \frac{1}{x^2} = \frac{1}{x^2} \left(1 - \frac{1}{1 + \frac{1}{x}}\right) > 0 \end{aligned}$$

For $x = 1$: $g'(x) = \ln(2) - 1 < 0$, and for $x \rightarrow \infty$: $\lim_{x \rightarrow \infty} \left(\ln\left(1 + \frac{1}{x}\right) - \frac{1}{x}\right) = 0$. Since $g'(x)$ is strictly increasing in x , we have that it must be the case that $g'(x) < 0 \forall x$. Hence, the minimum of $g(x)$ is reached at $x \rightarrow \infty$, in which case it is:

$$\lim_{x \rightarrow \infty} (x+1) \ln\left(1 + \frac{1}{x}\right) = 1.$$

Combining:

$$\begin{aligned} &(k-1) (\ln(k) - \ln(k-1)) - (N-k+1) (\ln(N-k+1) - \ln(N-k)) \\ &= (k-1) \ln\left(1 + \frac{1}{k-1}\right) - (N-k+1) \ln\left(1 + \frac{1}{N-k}\right) \\ &< \max_x \left\{x \ln\left(1 + \frac{1}{x}\right)\right\} - \min_x \left\{(x+1) \ln\left(1 + \frac{1}{x}\right)\right\} = 1 - 1 = 0. \end{aligned}$$

Therefore, $\frac{U^*(k+1)}{U^*(k)} < 1$ for $\alpha \rightarrow 1$, and by continuity for all α sufficiently close to one. Since $U^*(k)$ is strictly decreasing in k in this case, $k = 1$ is optimal.

Case 2: $N\mathbb{E}[\theta] < I$. In this case, the first term in (10) is negative. We first show that the equilibrium probability of the affirmative vote in the champions model, $p_1^* \equiv \arg \max_p \bar{U}(k, p)$, approaches zero, i.e., the same value as in the zero-NPV case. Denoting $NPV_0 \equiv N\mathbb{E}[\theta] - I$,

Pareto distribution implies $Nx_m = (I + NPV_0) \frac{\alpha-1}{\alpha}$. Thus,

$$\begin{aligned} p_1^* &= \arg \max_p NPV_0 \left(1 - (1-p)^{N-1} \right) + (1-p)^{N-1} \left(N \frac{\alpha}{\alpha-1} x_m p^{1-\frac{1}{\alpha}} - Ip \right) \\ &= \arg \max_p (1-p)^{N-1} \left(N \frac{\alpha}{\alpha-1} x_m p^{1-\frac{1}{\alpha}} - Ip - NPV_0 \right) \\ &= \arg \max_p (1-p)^{N-1} \left((I + NPV_0) p^{1-\frac{1}{\alpha}} - Ip - NPV_0 \right) \end{aligned}$$

When $\alpha \rightarrow 1$, $p^{1-\frac{1}{\alpha}} \rightarrow 1$ for any $p \in (0, 1)$, so

$$p_1^* \rightarrow \arg \max_p (1-p)^N = 0.$$

Next, re-write (8) as

$$U^*(k) = U_1^*(k) + U_2^*(k),$$

where

$$\begin{aligned} U_1^*(k) &= NPV_0 \left(1 - I_{1-p_k^*}(N-k, k) \right), \\ U_2^*(k) &= C_{N-1}^{k-1} (p_k^*)^{k-1} (1-p_k^*)^{N-k} \left(\int_{p_k^*}^1 (Nu(z) - I) dz \right), \end{aligned}$$

where $p_k^* \equiv \arg \max_p \bar{U}(k, p)$. We will show that in the limit $\alpha \rightarrow 1$, $U_2^*(1) > U_2^*(k)$ and $U_1^*(1) > U_1^*(k) \forall k > 1$.

The proof of $U_2^*(1) > U_2^*(k) \forall k > 1$ is as follows:

$$\begin{aligned} \lim_{\alpha \rightarrow 1} U_2^*(1) &= \lim_{\alpha \rightarrow 1} \max_p (1-p)^{N-1} \left(\frac{\alpha}{\alpha-1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha-1} p \right) Nx_m \\ &> \lim_{\alpha \rightarrow 1} \max_p C_{N-1}^{k-1} p^{k-1} (1-p)^{N-k} \left(\frac{\alpha}{\alpha-1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha-1} p \right) Nx_m \\ &\geq \lim_{\alpha \rightarrow 1} U_2^*(k) \end{aligned}$$

for all $k > 1$. Here, the first equality is from the fact that $p_1^* \rightarrow 0$ for both the negative-NPV

case and the zero-NPV case; the first inequality is from the fact that

$$\lim_{\alpha \rightarrow 1} \max_p C_{N-1}^{k-1} \max_p p^{k-1} (1-p)^{N-k} \left(\frac{\alpha}{\alpha-1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha-1} p \right) N x_m$$

is strictly decreasing in k , established in the proof of the zero-NPV case; and the second inequality is from the fact that $U_2^*(k)$ cannot exceed the point at which

$$C_{N-1}^{k-1} p^{k-1} (1-p)^{N-k} \left(\frac{\alpha}{\alpha-1} p^{1-\frac{1}{\alpha}} - \frac{\alpha}{\alpha-1} p \right) N x_m$$

is maximized over p .

The proof of $U_1^*(1) > U_1^*(k) \forall k > 1$ is as follows. Note that $\lim_{\alpha \rightarrow 1} U_1^*(1) = 0$, since $\lim_{\alpha \rightarrow 1} p_1^* = 0$. In contrast, for any $k > 1$, $\lim_{\alpha \rightarrow 1} U_1^*(k) < 0$, since $NPV_0 < 0$ and $p_k^* > 0$.

Since $U_2^*(1) > U_2^*(k)$ and $U_1^*(1) > U_1^*(k) \forall k > 1$, $U^*(1) > U^*(k)$ for any $k > 1$. Therefore, $k = 1$ is optimal.

Proof of Proposition 4

Consider the problem of a planner who observes N out of M signals (all signals observed by N partners) and decides on the investment decision. The expected value of investment is

$$\mathbb{E} [\min \{ \theta_1, \dots, \theta_M \} | \theta_1, \dots, \theta_N] = \mathbb{E} \left[\min \left\{ \tilde{\theta}_{(1:N)}, \tilde{\theta}_{(N+1:M)} \right\} | \tilde{\theta}_{(1:N)} \right],$$

where $\tilde{\theta}_{(A:B)} \equiv \min \{ \theta_A, \theta_{A+1}, \dots, \theta_B \}$. The equality follows from independence of all characteristics. Notice that this expression is a strictly increasing function of $\tilde{\theta}_{(1:N)}$: this follows from $\min \left\{ \tilde{\theta}_{(1:N)}, \tilde{\theta}_{(N+1:M)} \right\}$ being weakly increasing in $\tilde{\theta}_{(1:N)}$ and $\tilde{\theta}_{(N+1:M)}$ taking values below $\tilde{\theta}_{(1:N)}$ and above $\tilde{\theta}_{(1:N)}$ with positive probabilities for any finite $\tilde{\theta}_{(1:N)} > -l$. Therefore, the optimal decision of the planner is to invest in the project if and only if $\tilde{\theta}_{(1:N)} \geq \tilde{\theta}^*$, where threshold $\tilde{\theta}^*$ is given by

$$\mathbb{E} \left[\min \left\{ \tilde{\theta}^*, \tilde{\theta}_{(N+1:M)} \right\} \right] = I.$$

Next, we show that the voting rule $k = N$ implements this optimal decision of the planner. To do this, we show that voting threshold $\hat{\theta} = \tilde{\theta}^*$ corresponds to a Nash equilibrium in the voting game when $k = N$. Consider member i with signal θ_i . Her vote is pivotal if and only if all $N - 1$ committee members vote for the project, which implies that all of their signals exceed $\hat{\theta}$. Consider member i 's forecast of V , conditional on her signal θ_i and on the information that her vote is pivotal. If $\theta_i \leq \hat{\theta}$, then

$$\begin{aligned}\mathbb{E}[V|\theta_i, Piv_i] &= \mathbb{E}\left[\min\left\{\tilde{\theta}_{(1:N)}, \tilde{\theta}_{(N+1:M)}\right\} \mid \theta_i, \theta_j \geq \hat{\theta} \ \forall j \neq i, j \leq N\right] \\ &= \mathbb{E}\left[\min\left\{\theta_i, \tilde{\theta}_{(N+1:M)}\right\}\right] \leq \mathbb{E}\left[\min\left\{\hat{\theta}, \tilde{\theta}_{(N+1:M)}\right\}\right].\end{aligned}$$

The second equality follows from $\min\left\{\tilde{\theta}_{(1:N)}\right\} = \theta_i$, when $\theta_i \leq \hat{\theta}$ and member i is pivotal. The inequality follows from weak monotonicity of the minimum function in θ_i and the fact that $\theta_i \leq \hat{\theta}$. If $\theta_i \geq \hat{\theta}$, then

$$\begin{aligned}\mathbb{E}[V|\theta_i, Piv_i] &= \mathbb{E}\left[\min\left\{\tilde{\theta}_{(1:N)}, \tilde{\theta}_{(N+1:M)}\right\} \mid \theta_i, \theta_j \geq \hat{\theta} \ \forall j \leq N\right] \\ &\geq \mathbb{E}\left[\min\left\{\tilde{\theta}_{(1:N)}, \tilde{\theta}_{(N+1:M)}\right\} \mid \theta_i = \hat{\theta}, \theta_j \geq \hat{\theta} \ \forall j \leq N\right] \\ &= \mathbb{E}\left[\min\left\{\hat{\theta}, \tilde{\theta}_{(N+1:M)}\right\}\right]\end{aligned}$$

The inequality follows from weak monotonicity of the minimum function in θ_i and the fact that $\theta_i \geq \hat{\theta}$. The equality follows from $\tilde{\theta}_{(1:N)} = \hat{\theta}$ when $\theta_i = \hat{\theta}$ and $\theta_j \geq \hat{\theta} \ \forall j \in \{1, \dots, N\}$. We can see that the strategy profile with voting threshold $\hat{\theta} = \tilde{\theta}^*$ constitutes a Nash equilibrium. Indeed, if $\hat{\theta} = \tilde{\theta}^*$ and $\theta_i \leq \tilde{\theta}^*$, member i is better off voting against the project since $\mathbb{E}[V|\theta_i, Piv_i] \leq \mathbb{E}\left[\min\left\{\tilde{\theta}^*, \tilde{\theta}_{(N+1:M)}\right\}\right] = I$; if $\hat{\theta} = \tilde{\theta}^*$ and $\theta_i \geq \tilde{\theta}^*$, member i is better off voting for the project since $\mathbb{E}[V|\theta_i, Piv_i] \geq \mathbb{E}\left[\min\left\{\tilde{\theta}^*, \tilde{\theta}_{(N+1:M)}\right\}\right] = I$. Since the consensus voting rule implements the planner's optimal decision, no other voting rule can improve on it. Therefore $k = N$ must be optimal.

C Extensions

C.1 Model with Heterogeneous Preferences

Given the apparent value to investors using the champions models in such contexts, it seems interesting that this practice is observed in VC but has not been noted in other contexts such as corporate R&D. In this extension of the model, we provide one potential explanation: the fact that the champions model loses its value if each committee member may prefer to vote for the project for a private, rather than common, reason.

Consider the following variation of the baseline model. Suppose that with probability π each partner wants to either do the project or not do a project (with equal probabilities) for a private reason irrespective of its quality (i.e., the project is a private benefit or private cost project). This may be due to agency conflicts or due to career concerns in the organization. In contrast, with probability $1 - \pi$, the partner is unbiased and wants to maximize the expected value. Suppose that the realizations of whether the project is partner i 's private benefit project are independent across partners.

On one extreme, if $\pi = 0$, then all partners have common objectives of maximizing the value of the project - which reverts to the model of the previous section. On the other extreme, if $\pi = 1$, then all partners have private values.

Under this model, each partner will vote for the project in two scenarios. First, she will vote for the project with probability $\frac{1}{2}\pi$, irrespective of the realization of her signal. Second, with probability $1 - \pi$, she will vote for the project if and only if it exceeds some cut-off $\hat{\theta}_k(\pi)$. Cut-off $\hat{\theta}_k$ is implicitly defined by the following equation:

$$\hat{\theta}_k + ((k - 1) \mathbb{E}[\theta_i | i \text{ votes for}] + (N - k) \mathbb{E}[\theta_i | i \text{ votes against}]) = I \quad (11)$$

To see the role of private benefits, consider what happens in the extreme case of $\pi \rightarrow 1$. In this case, the vote of each other partner is determined fully by private benefits (or costs) of

the project. Thus, these votes are uninformative about the partners' signals, so the equilibrium cut-off of the benevolent partner will be $\hat{\theta}_k = I - (N - 1) \mathbb{E}[\theta]$. In particular, her voting strategy is independent of the rule k . Consider $k = 1$. Compared to the case of $\pi = 0$, each partner will now overchampion for the project for two reasons. First, she will overchampion if she wants to do the project for a private benefit reason, since the probability of approval vote in this case ($\frac{1}{2}$) exceeds that of p_1 in Proposition 3 (below $\frac{1}{N}$ for any α). Second, and more interestingly, she will overchampion even if she is benevolent: This will happen because the negative votes of others do not reveal negative information about other characteristics of the project when the votes occur primarily for private benefit reasons.

It is easy to see why the champions model will be suboptimal for a sufficiently high π in this case. While the champions model is still useful at identifying tail projects if the champion happens to be benevolent, most projects get championed for private benefit reasons. Since ex-ante an average project has negative NPV, it is optimal to increase k to reduce the probability of investment for private benefit reasons. Decision making rules that require more consensus among committee members are better in this case, because they curb overchampioning.

C.2 Fat Left Tails

The model can be symmetrically applied to study investments in projects with fat left tails. We can turn the parameters of the model of the previous sections on their head to analyze this question. Suppose now that, as before, the valuation of the company is as before: $V = \theta_1 + \theta_2 + \dots + \theta_M$, but now suppose that each of the M characteristics is distributed independently over $[-\infty, r)$ (where $r > 0$) according to a distribution function $H(\cdot)$ with density $h(\cdot)$. Further assume, as before that G and g are such that $E[\theta_i] = 0$ for all i . This distribution is meant to capture the shape of late stage investment decisions (or, alternatively, the payoff profile of debt contracts as opposed to equity contracts). Under this profile, the conclusions of the previous section reverse:

Proposition 5. *Suppose that H is subexponential. Then, the unanimity rule rejects a project with $V \rightarrow -\infty$ with probability one. In contrast, any $k < N$ rule accepts such a project with a strictly positive probability.*

The intuition of this proposition is very similar to that of proposition 1: unanimity and champions rules are two sides of the same coin. Under unanimity only one partner needs to object to the project to reject it; under the champions rule only one partner needs to support the project to accept it.

C.3 Model with Noisy Signals of Partners

In the main model, we assume that partners get perfectly precise signals about attributes of the project. In this extension, we show that this assumption is not critical, and results continue to hold when signals of partners contain noise.

Suppose that partner i observes signal S_i distributed with density $g(S_i)$, which is informative about θ_i , but uninformative about any other characteristic of the project. Let $h(\theta_i|S_i)$ denote density of characteristic θ_i , conditional on partner i 's signal being S_i . We assume that (S_i, θ_i) satisfy the strict monotone likelihood ratio property: likelihood ratio $\frac{h(\theta_i|S_i)}{h(\theta_i|S'_i)}$ is strictly increasing in θ_i for any S_i and $S'_i < S_i$. Intuitively, higher realizations of signal S_i are more indicative of higher values of characteristic θ_i . We also normalize, without loss of generality, the conditional expected value of θ_i to be equal to S_i : $\int \theta_i h(\theta_i|S_i) d\theta_i = S_i$.²⁰ Intuitively, signal S_i is defined as the best guess of partner i about characteristic θ_i .

Example. As an example of this setting, consider $\theta_i = A_i S_i$, where S_i follows Pareto distribution with shape α and scale x_m , and A_i is lognormal with a mean of one, independent from all other random variables of the model. Intuitively, S_i captures the best guess of partner i about the value of her characteristic, while A_i captures the possibility that partner i is wrong (the greater the variance of A_i the less precise signal S_i is).

²⁰See, e.g., [DeMarzo, Kremer, and Skrzypacz \(2005\)](#) for a similar normalization.

From this point, the analysis repeats the analysis of Section 3.3 with the change that we now use the distribution of signals as a primitive rather than the distribution of characteristics. In particular, if the voting protocol requires k or more positive votes for investment, then each partner votes for investment if and only if her signal exceeds cutoff \hat{S}_k , given by

$$\hat{S}_k + (k - 1) \mathbb{E} \left[S | S \geq \hat{S}_k \right] + (N - k) \mathbb{E} \left[S | S \leq \hat{S}_k \right] + (M - N) \mathbb{E} [\theta] = I.$$

Since the model is equivalent to the model of Section 3.3, variants of Propositions 1-3 in which we substitute distribution of signals for distribution of characteristics, also hold. The proofs of these propositions are identical to the proofs in the main model.

D Survey Questions and Quotations

D.1 Survey Questions

3. Please indicate the formal process that best describes your investments in (a) seed stage, (b) early stage, and (c) later/growth stage.

	Seed	Early stage	Later/growth stage
Not applicable (don't make this kind of investment)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Require unanimous agreement to do the deal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Require majority of votes to do the deal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Require majority of votes as long as no veto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A single partner can do the deal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A single partner can do the detail as long as no veto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Leaving aside your formal investment process, what is the culture you feel best describes your partnership?

	Seed	Early stage	Later/growth stage
Not applicable (don't make this kind of investment)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
An enthusiastic champion is usually sufficient for others to vote 'yes'	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Majority of the partners have to be enthusiastic to do the deal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Only deals where <i>all</i> partners are enthusiastic get done	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please describe below)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you selected "other" above, please describe in the box below.

Figure 12: Key Survey Questions used in Analysis: This figure is a screenshot of two of the survey questions asked to VC investors about the formal and informal voting processes their firms use to make decisions.

3. A prior study has documented that many VC partnerships use a champions voting rule where a single partner can decide whether or not to do a deal, even if the majority of partners is not supportive. These VCs often use champions voting *for seed and early investments but not for later-stage investments*. How much do you agree with the following possible explanations?

Expertise: VCs use champions voting for seed and early investments but not for later-stage investments because earlier-stage deals require more specialized expertise than late-stage deals, so partnerships give people with expertise greater discretion in doing earlier-stage deals but not later-stage deals.

Strongly disagree <input type="radio"/>	Somewhat disagree <input type="radio"/>	Neither agree nor disagree <input type="radio"/>	Somewhat agree <input type="radio"/>	Strongly agree <input type="radio"/>
--------------------------------------------	--------------------------------------------	-----------------------------------------------------	-----------------------------------------	-----------------------------------------

Catching Outliers: VCs use champions voting for seed and early investments but not for later-stage investments because the best early-stage investments can be extremely promising in some dimensions but flawed in others. Requiring too much consensus can therefore lead partnerships to pass on the best deals.

Strongly disagree <input type="radio"/>	Somewhat disagree <input type="radio"/>	Neither agree nor disagree <input type="radio"/>	Somewhat agree <input type="radio"/>	Strongly agree <input type="radio"/>
--------------------------------------------	--------------------------------------------	-----------------------------------------------------	-----------------------------------------	-----------------------------------------

Effort: VCs use champions voting for seed and early investments but not for later-stage investments because they care less about smaller deals, and don't want too many partners to spend effort on diligence. Leaving the decision to a champion for these smaller deals also incentivizes greater effort on the part of the champion.

Strongly disagree <input type="radio"/>	Somewhat disagree <input type="radio"/>	Neither agree nor disagree <input type="radio"/>	Somewhat agree <input type="radio"/>	Strongly agree <input type="radio"/>
--------------------------------------------	--------------------------------------------	-----------------------------------------------------	-----------------------------------------	-----------------------------------------

Figure 13: Key Survey Questions used in Analysis: This figure is a screenshot of the survey question asked to VC investors about the reasons for champions voting processes their firms use to make decisions.

D.2 Quotations

Here we provide some quotations from the anonymous participants in our surveys:

We use champion voting. The primary reason is that successful early-stage VC requires people to "think differently" from the pack (either by being early to a trend or literally interpreting the same fact set differently). Requiring consensus risks cautious investments that regress to the mean. [One investment] is a great example of a "Think Differently" investment that yielded a big return for us but was unpopular in the partnership initially.

At least in our firm we recognize that most of our alpha comes from early-stage conviction, but that this conviction requires often a deep up close study of a trend or prospect, and that knowledge can be hard to communicate to a larger group. So we leave early stage dealmakers alone on early stage decisions worried that if we provide too much feedback they'll lose that conviction on the deals that matter. It should also be noted that some of the best investments are both weird/different and can be at valuations that don't always completely compensate for this weirdness (because there's momentum) - and so the early stage conviction of the champions can be fragile if they hear lots of pushback from their partners. We have many examples of an investor building conviction early in their career on a future potential fund moving deal only to have that conviction shattered by an errant comment by an "old dog" investor.

It's mostly because in early stages neither the trend nor the metrics are obvious and committee decisions can result in the firm losing out on outliers

We care more about early stage deals since we can get more ownership at a lower price, but have a partner conviction model since our partners each have different areas of expertise. Our partner conviction model also stands for later stage deals, but more people are involved in diligence to help the partner make a better decision.

In practice, there is 1-2 partners who are responsible for ongoing support for the portfolio. It is more important to have strong support from 1 partner, vs. lukewarm support from multiple partners. In the latter case no one will truly assume responsibility for post investment support.