Organizational decision making: An information aggregation view[1]

Felipe A. Csaszar[a], J. P. Eggers[b]
[a]Ross School of Business, U. of Michigan, 701 Tappan Ave., Ann Arbor, MI 48109
Tel: (734) 615 4854, Email: fcsaszar@umich.edu
[b]New York University, Stern School of Business, New York, NY 10012
Tel: (212) 998 0874, Email: jeggers@stern.nyu.edu

December 3, 2012

Forthcoming in Management Science

[1]For helpful comments the authors thank Dan Levinthal, Tony Marciano, Gabriel Natividad, Narayan Pant, David Ross, José Santos, Gus Stuart, and Timothy Van Zandt as well as seminar participants at INSEAD, New York University, the University of Michigan, the Tuck School of Business at Dartmouth, the Washington University in St. Louis, the European School of Management & Technology, the TOM Society Meeting at Barcelona, the Atlanta Competitive Advantage Conference, the Academy of Management Conference, and the Utah-BYU Winter Conference. The authors thank Naveen Naidu Narisetty for his contributions as a research assistant. All errors remain our own.

Abstract

We study four information aggregation structures commonly used by organizations to evaluate opportunities: individual decision making, delegation to experts, majority voting, and averaging of opinions. Using a formal mathematical model, we investigate how the performance of each of these structures is contingent upon the breadth of knowledge within the firm and changes in the environment. Our model builds on work done in the Carnegie tradition and in the group and behavioral decision making literatures. We use the model to explore when delegation is preferable to other structures, such as voting and averaging. Our model shows that delegation is the most effective structure when there is diversity of expertise, when accurate delegation is possible, and when there is a good fit between the firm's knowledge and the knowledge required by the environment. Otherwise, depending on the knowledge breadth of the firm, voting or averaging may be the most effective structure. Finally, we use our model to shed light on which structures are more robust to radical environmental change and when crowd-based decision-making may outperform delegation.

Keywords: organizational structure, decision making, knowledge, environmental change

1 Introduction

Important organizational decisions are likely to be made by groups rather than by single individuals (Tindale et al. 2003:381). Such groups might include top management teams, boards of directors, or finance committees. Given that different group members may often possess different opinions, how individual opinions are aggregated into a group-level decision has important implications for organizational action and performance. But selecting the right method to aggregate opinions can be challenging, as there are many aggregation methods and which one is the best may be contingent on multiple factors. The following example illustrates some of these challenges.

Imagine the case of three founders of a startup company deciding on whether or not to acquire a competitor. How should their different opinions on the value of the target be aggregated? Should they vote on whether to acquire or not, should they delegate the decision to one member, or should they average their differing valuations and make a decision based on that average? Would the answer depend on the expertise of the decision makers (e.g., if they were three MBAs versus if they were an accountant, an MBA, and a scientist)? Would the answer depend on the degree of environmental uncertainty (as when, e.g., the target produces a promising but unproved technology versus a commodity)?

Problems such as these are common in settings ranging from small startups to large governments. Example settings include movie studio executives or corporate finance committees deciding on which projects to fund, venture capitalists and mutual fund managers choosing which assets to buy, and boards of directors and top management teams deciding on strategic actions. In general, information aggregation pervades organizational decision making.

Despite this pervasiveness in organizations, with few exceptions (e.g., Knudsen and Levinthal 2007, Fang et al. 2010, Csaszar 2012), the organizations literature has devoted comparatively little attention to information aggregation issues. Information aggregation, which is also known as the *structure* of decision making, was proclaimed by the Carnegie tradition as one of its central concerns (Cyert and March 1963:19–22). In fact, Simon (1947/1997:18–19) defines the basic construct of "organization" in terms of information aggregation: "the pattern of communications and relations among a group of human beings, including the processes for making and implementing decisions." Despite this deep concern with structure and information aggregation, a recent article by Gavetti, Levinthal, and Ocasio (2007:528) notes that this is one of the "forgotten pillars" of the Carnegie tradition. Other literatures that share an information processing sensibility—such as the resource allocation process (Bower 1970) and organization design (Thompson 1967, Burton and Obel 2004)—have also discussed the lack of research on the topic (Miller et al. 2009).

Our central question is this: What is the most appropriate decision-making structure for the evaluation of potential alternatives given an organization's environment and the expertise of its members? To address this question in a rigorous manner, we develop a parsimonious model of information aggregation in organizations, which allows us to investigate questions of organizational design in a manner similar to recent work by Knudsen and Levinthal (2007) and Csaszar (2013).

Our modeling approach draws on and extends models of information aggregation from psychology and economics. Previous models in these literatures have analyzed (among other issues) the robustness of different structures (Hastie and Kameda 2005), optimal processes for weighting opinions (Ben-Yashar and Nitzan 1997), and when simple heuristics are capable of making near-optimal decisions (Gigerenzer and Goldstein 1996). We contribute to these literatures by introducing to existing information aggregation models a few key organizational elements. These include assessing the robustness of our focal decision-making structures—delegation to experts, majority voting, and averaging of opinions—to changes in the environment and the expertise of its members. Additional contingencies include errors in delegation, errors in assessing expertise, and updating processes by decision makers. Our approach captures the knowledge possessed by the firm and the knowledge required by the environment through modeling individuals' perceptions, knowledge, and the environment as a stochastic process. This approach allows us to extend models in behavioral and group decision making, leading to novel results that are organizational in character. Our model also provides a starting point for investigating further questions of aggregation in organizations.

Three novel findings emerge from this analysis. First, the interconnected nature of structure, knowledge, and environment means that each decision-making structure can be the optimal choice under the appropriate set of contingencies. Second, organizational performance depends in a non-trivial way on how the opinions of individual members are aggregated. This is true even for means of aggregation that are ostensibly similar, such as averaging estimates and allowing individuals to vote. Third, the delegation of decision-making authority can be surprisingly effective, specifically when the individual decision makers possess significantly different knowledge, when projects can be correctly assigned to suitable individuals, and when there is a good fit between the knowledge of the firm and the knowledge required by the environment. These findings provide insights into the prevalence of delegation as a means of organizational decision-making; the conditions under which other, more egalitarian structures can be effective within a firm; and the relative performance of different information aggregation mechanisms available to organizations.

2 Theoretical Motivation

The main question explored in this paper ("What is the most appropriate decision-making structure for the evaluation of potential alternatives given an organization's environment and the expertise of its members?") has parallels with a number of questions that have been explored in different literatures interested in information aggregation. In this section we review these literatures and discuss how our approach differs from previous studies.

2.1 Antecedents of Our Work

The study of different information aggregation rules has a long history. In fact, the study of the main rules explored in this paper—delegation, voting, and averaging—can be traced back at least to Aristotle (c.330 BCE/1984), Condorcet (1785/1994), and Laplace (1814/1995), respectively.[1] Because devising effective aggregation rules is relevant in many applications, research on this topic is vast and spans multiple disciplines (Grofman and Owen 1986:xi). Therefore, the following review is focused on prior research with close similarity to the current study. Thus, this review is biased toward models of aggregation rules that are carried out by fallible decision makers who share a common goal. We group these literatures into four clusters: Carnegie tradition, behavioral decision making, group decision making, and economics.

Carnegie Tradition. The Carnegie tradition is concerned with how decisions are made within organizations. Thus, as mentioned in the Introduction, issues of information aggregation are central to this tradition. In recent years, several models of organizational decision making have emerged from this tradition, very much in the spirit of classic work (i.e., Cohen et al. 1972, March 1991).

To understand the previous research and how the current work relates to it, it is useful to look at the three main stages of organizational decision making: (i) search, (ii) evaluation, and (iii) implementation of alternatives (Mintzberg et al. 1976, Schwenk 1984). Interestingly, most of the models of organizational decision making done in the Carnegie tradition have focused on stage (i) by using NK and other simulation methods (e.g., Seshadri and Shapira 2003, Rivkin and Siggelkow 2003, Siggelkow and Rivkin 2005).

This paper models how a group of decision-makers evaluate alternatives, and thus falls squarely into the less-researched stage (ii) of the previous taxonomy. Among the few studies in the Carnegie tradition that have modeled aspects of the evaluation stage, there is work about the effect of different organizational structures on errors of omission and commission (Knudsen and Levinthal 2007, Christensen and Knudsen 2010, Csaszar 2012) and on exploration and exploitation (Fang et al. 2010, Csaszar 2013).

Behavioral Decision Making. This literature is concerned with the performance of different decision-making heuristics. A prevalent finding in this literature is that several heuristics, that is, rules that do not take into account all of the available data or whose internal parameters do not stem from an optimization process, can perform at similar levels compared with optimal decision rules in many realistic settings. Some of the heuristics that have been shown to exhibit this high performance are unit weights (Einhorn and Hogarth 1975), improper linear models (Dawes 1979), and take-the-best (Gigerenzer and Goldstein 1996) and similar lexicographic rules (e.g., Deterministic Elimination by Aspects; Hogarth and Karelaia 2005).

As do the structures in our model, the previous heuristics aim at choosing the best among a set of alternatives, each alternative being described by a set of characteristics or cues. If one interprets the judgments of the individuals in our model as decision-making cues, two similarities emerge between our model and the aforementioned heuristics. First, averaging of judgments is akin to unit

[1] Aristotle observed that groups can outperform individuals because, in a group, "some understand one part, and some another, and among them they understand the whole" (Aristotle, c.330 BCE/1984, book III.11), implicitly pointing to a delegation process inside the group. Centuries later, Condorcet (1785/1994) proved that majority voting would tend to select the right choice as the number of voters approached infinity, assuming uncorrelated voters with a modicum of screening ability. A few years later, Laplace (1814/1995, chapter IX) suggested that averaging individual opinions could be used to synthesize a highly accurate opinion.

weights (i.e., the judgments of the experts are added up using equal weights). Second, delegation to the most suitable expert is akin to take-the-best, as only the most relevant cue is used (i.e., the judgment of the most suitable expert). A more general similarity relates to one of the goals of our model: to understand what decision rules work better under what environments. This goal was highlighted early in the behavioral decision literature (Brunswik 1952, Simon 1956) and has more recently motivated a rich stream of literature on adaptive decision making (Payne et al. 1993, Hogarth and Karelaia 2007, Soll and Larrick 2009).

Group Decision Making. The group decision making literature has modeled different rules that groups can use to aggregate opinions. For instance, Hastie and Kameda (2005), in a similar vein to earlier work by Einhorn et al. (1977) and Hogarth (1978), study the performance of nine aggregation rules. The main aggregation rules studied in the current paper—delegation, voting, and averaging—are among the ones explored by Hastie and Kameda (2005) (they call these rules best-member, majority/plurality, and average winner, respectively). They test the performance of these decision-making rules under an environment defined by a linear combination of the cue values. Under this environment, averaging is the best performer, as it naturally captures the linearity of the environment. They also find that voting performs similarly to averaging and that voting usually outperforms delegation. In general, they find that rules that combine individual scores (such as averaging) tend to outperform rules that do not combine individual scores (such as delegation and voting). More recently, Kameda et al. (2011) have shown that majority voting can pool information effectively even if the group members are self interested and some of them decide to shirk their duty to vote.

Soll and Larrick (2009) also develop a model that explores some of the same structures as the current paper. In particular, they use their model to provide a normative framework to choose between averaging and delegation, and then perform experiments to determine how people deviate from this framework. Two of the parameters in their framework (dispersion in judges' accuracies and identifiability of the most expert judge) are similar to elements in our model (namely, the relative accuracy of the experts and delegation errors, which are introduced in sections 4.1 and 4.3, respectively). Larrick and Soll (2006) show that, in two-member groups, averaging performs at least as well as delegating to a random member, but despite this, people often distrust averaging. Newell and Shanks (2003) use experiments to show when people rely on a single cue versus multiple cues, and thus speak to the issue of when people use delegation versus rules that use more information. For surveys of the group decision making literature, see Tindale et al. (2003) and Laughlin (2011).

Economics. Economics has also studied models of information aggregation. Research here has studied, among other issues, the efficiency of decentralized, hierarchical structures (Radner 1993); compared hierarchy and polyarchy as means of aggregating dichotomous choices (Sah and Stiglitz 1986); and devised optimal weights for committee decision making (Ben-Yashar and Nitzan 1997).[2]

[2] A weighted averaging rule is optimal given an optimal set of weights. But computing these weights requires of assumptions that are not particularly plausible from a behavioral standpoint. In Section 4.4 we discuss these assumptions and use the optimal rule as a benchmark against which to compare the performance of the structures studied in this paper.

Voting models are also common in social choice theory, however, these models are of a different nature than the information aggregation models discussed in this paper: in the models we have reviewed, the individuals share a utility function (e.g., they are all trying to choose what is the best choice for a given firm), thus differences of opinion arise because of errors in perception or reasoning. In contrast, models of social choice assume that different individuals have different utility functions (e.g., individuals have different preferences and like different political candidates), which leads to the well-known impossibility theorem (Arrow 1951), that is, to the inability of aggregating individual preferences into group-level preferences that satisfy a reasonable set of fairness criteria.

2.2 How This Study Differs From Previous Studies

Although the reviewed literatures have provided many valuable insights on the properties of different aggregation methods, these insights are difficult to apply directly to organizational settings, because organizational performance depends heavily on two contingencies—individuals' knowledge and environmental change—which have not been accounted for in previous information aggregation models. This study extends previous information aggregation models by infusing these organizational contingencies. The rationale for including both of these contingencies is discussed below.

First, our model considers individuals who are heterogeneous with respect to their *knowledge* or expertise. In contrast, previous group decision making research has conceptualized heterogeneity in terms of *ability* (Grofman et al. 1983, Kanazawa 1998, Sorkin et al. 2001, Soll and Larrick 2009), where a high ability agent produces, on average, superior estimates of the project under consideration than low ability agents. On the other hand, differences in knowledge imply that a given individual may be inferior to another at assessing Project A, but superior at assessing Project B. In the context of organizations, differences in knowledge are fundamental, as specialization and division of labor typically lead to decision-making teams involving members with different expertises (Bower 1970, Hambrick and Mason 1984). For instance, top management teams often involve individuals with financial, marketing, and operations expertise, who are not equally capable of assessing the quality of, say, a new manufacturing process.

Second, our model considers decision-making in both stable and changing environments. In contrast, previous group decision models have usually studied stable environments. In fact, Hastie and Kameda (2005:506) state that, "how various group decision rules will perform in unstable environments is an open question." This lacuna may be a result of the fact that the organizations typically studied in group decision making are formed "on demand" to solve a given problem in a given environment and thus need not persist, as firms do, over longer periods of time that entail changes in the environment. But in the context of organizations, environmental change plays a central role. This is evidenced by the vast and diverse literatures on topics such as disruptive technological change (Tushman and Anderson 1986), organizational ecology (Hannan and Freeman 1977), and strategic fit (Siggelkow 2001).

Including both contingencies allows our model to shed light on when it is better to use which structure in terms of contingencies that are central to organizations. It also leads to a number of new results regarding the relative advantages of each structure. For instance, Hastie and Kameda (2005) found that averaging outperformed voting, and voting outperformed delegation. In our model, each one of these three structures can be the best performer, depending on the value of the contingencies we study. To further understand decision making in organizations, later in the paper we incorporate additional contingencies into the analysis (such as changing the number of decision makers, including correlation among the individuals, and allowing for delegation errors). In sum, our approach extends previous models of information aggregation to account for distinctly organizational characteristics. We believe that introducing these elements allows our model to address important questions faced by organizations.

3 Model

The aim of our model is to study how organizational performance depends on organizational structure, the expertise of its members, and the organization's external environment. We start by providing a brief overview of the model before describing each of its components in detail.

The decision-making structures we analyze are in charge of screening a stream of projects—that is, approving good projects and rejecting the rest. Screening projects or opportunities is a common task in many settings, including formal organizations such as movie studio executives evaluating multi-million dollar projects (we discussed other examples in Section 1) and informal organizations such as mountaineers deciding on whether today is the best day for a summit attempt. The structures we study differ in how they aggregate the opinions of their members to produce organization-level decisions about the projects under review. For the sake of results comparability and model parsimony, most of our analyses compare three structures with three members each (we call these structures Delegation, Voting, and Averaging). For benchmarking purposes we also analyze an organization with a single decision maker (a structure that we call Individual). Later in the paper we also analyze variations of these structures and increase the number of decision makers.

The projects screened by the organizations in our model are described by a type and a quality. The project *type* represents the domain of knowledge, or expertise, involved in accurately assessing the project (e.g., in the venture capital context, project types could correspond to semiconductors, software, biotechnology, etc.). The project *quality* represents how much value the project will create if implemented. Project quality is noisily perceived by individuals, and the greater the distance between the individual's expertise and the type of project being evaluated, the greater the noise in that individual's perception of the project's quality (e.g., a software expert will be more accurate at determining the value of a software startup than of a semiconductor startup). Organizational *performance* is defined as the sum of the qualities of the projects approved by the firm, divided by the total number of projects considered.

3.1 Projects

We define a *project* as a tuple (q, t), where q denotes the project's quality and t its type. We interpret q as value, in terms of (appropriately discounted) revenues minus costs; thus we say a project is "good" only if $q > 0$. For simplicity, we assume that the discount rate is correctly set, that firms do not face liquidity constraints, and that there are no interactions among projects. Under these assumptions, firms maximize their performance by accepting good projects and rejecting bad ones. The type (t) of the project is a real number that denotes the specific type of knowledge required to assess the project's value properly. The actual value of t is relevant only with respect to the expertise of the individual decision makers, as discussed in Section 3.2.

We define the *project environment* as the range of projects that a firm faces. We view this environment as a set of problems exogenously posed to the organization, a view that is consistent with multiple research traditions (Hannan and Freeman 1977, Carley and Lin 1997). Thus, our model captures the evaluation of alternatives, not their generation (Knudsen and Levinthal 2007). Because projects are described by two parameters, the environment is defined by the ranges that both q and t can take. We assume that projects are uniformly distributed in the rectangular interval defined by $[\underline{q}, \overline{q}]$ and $[\underline{t}, \overline{t}]$.[3] Continuing with our example, these bounds could reflect the range of projects a venture capital firm faces in terms of quality (e.g., values from −\$100 million to \$100 million) and types (e.g., from 0 to 10, where 0 represents hardware projects, 10 represents software projects, and intermediate numbers represent mixtures between these two extremes).[4]

3.2 Individuals

As with projects, we model individuals as having a type, which we call an *expertise*. Heterogeneity in expertise (as discussed in Section 2.2) is largely absent from the group decision making literature, but is a common assumption in literatures on, for example, top management teams (Hambrick and Mason 1984) and organizational learning (Liang et al. 1995). For any given project, the difference between a project's type and the expertise of the individual assessing that project affects the level of noise in the individual's perception of project quality. For example, if a software expert is called upon to evaluate a project that involves mostly software, then her perception will likely be more accurate than the perception of a hardware expert. We model noisy perception as a signal plus noise, where the signal is the actual quality of the project (q) and the standard deviation of the noise is proportional to the distance between project type and individual expertise.[5] Mathematically, if an individual of expertise e is called to assess the quality of project (q, t), then she will perceive

[3] In Appendix B we consider an alternative probability distribution of project types.

[4] An alternative interpretation of $[\underline{t}, \overline{t}]$ is in terms of a filtering mechanism, i.e., the firm ignores the projects that fall outside that range.

[5] Our assumption that the project type and individual expertise are real numbers can be interpreted in a nonrestrictive manner (à la Hotelling 1929), as simply meaning that it is possible to compute a distance between a given project type and a given individual expertise. In other words, we make no assumptions about the dimensionality of the knowledge space and assume only that it is possible to compute distances in that space.

the quality of this project as

$$q' = q + \tilde{n}, \text{ where } \tilde{n} \sim N(0, |t - e|). \tag{1}$$

We thus denote the perceived quality as q'.[6] Modeling perception as signal plus noise is consistent with prior models (e.g., Einhorn et al. 1977, Knudsen and Levinthal 2007) and with empirical work on managerial perceptions (e.g., Mezias and Starbuck 2003).

3.3 Structures

In the spirit of Simon's (1947/1997:18–19) definition of organization as "the pattern of communications and relations among a group of human beings, including the processes for making and implementing decisions," we model organizational structure as the mechanism that aggregates individual perceptions into a group-level decision regarding each project reviewed.

While there are an unlimited number of potential decision-making structures to study, we focus on four in the main part of this paper because they are representative of structures commonly used in real-world organizations.[7] These four structures are graphically summarized in Figure 1. Each of these structures (except for the first) consists of three individuals with expertise levels e_L, e_M, and e_H such that $e_L \leq e_M \leq e_H$. To ease presentation of results and to simplify our discussion of the homogeneity or heterogeneity of expertise, we express the three expertise levels in terms of two parameters, e_M (the expertise of the intermediate individual) and β (the breadth of knowledge in the organization) such that $e_L = e_M - \beta$ and $e_H = e_M + \beta$. We discuss an alternative distribution of expertise in Appendix B. The following paragraphs describe how each of the four structures operates.

Individual (Figure 1a). This is the simplest structure; it involves only one manager, with expertise e_M, who is in charge of either accepting or rejecting projects. If he perceives a project to have a positive quality (i.e., $q' > 0$) then he approves the project; otherwise, he rejects it. We use this as a benchmark against which the other, more complex structures are compared. This structure is representative of settings where decisions are made by a single individual; examples include many small and entrepreneurial firms as well as firms with a particularly powerful manager. This structure can also be characterized as being highly centralized, as authority is closely held by a single decision maker (Huber and McDaniel 1986:581).

Delegation (Figure 1b). In this structure, projects are delegated to the manager whose expertise is closest to the type of the project being screened (e.g., if a project type $t = 4$ is assessed by a structure with experts $e_L = 0$, $e_M = 5$, and $e_H = 10$, then the project would be assigned to

[6]The noise term in equation (1) implies that if expertise equals project type then perception is completely accurate. One might argue that this standard of perfection is too high, even for an expert with $t = e$. It is possible to relax this assumption by adding a baseline error rate, so that no individual can achieve perfect perception. Doing so does not result in any qualitative change to our results, as we discuss in Section 4.4 and Appendix B.

[7]In Section 4.4 and Appendix C we discuss a number of other structures, namely structures in which individuals' errors are correlated, structures that optimally weight individuals' judgments, and structures whose individuals' update their judgments after taking into account their level of competence.

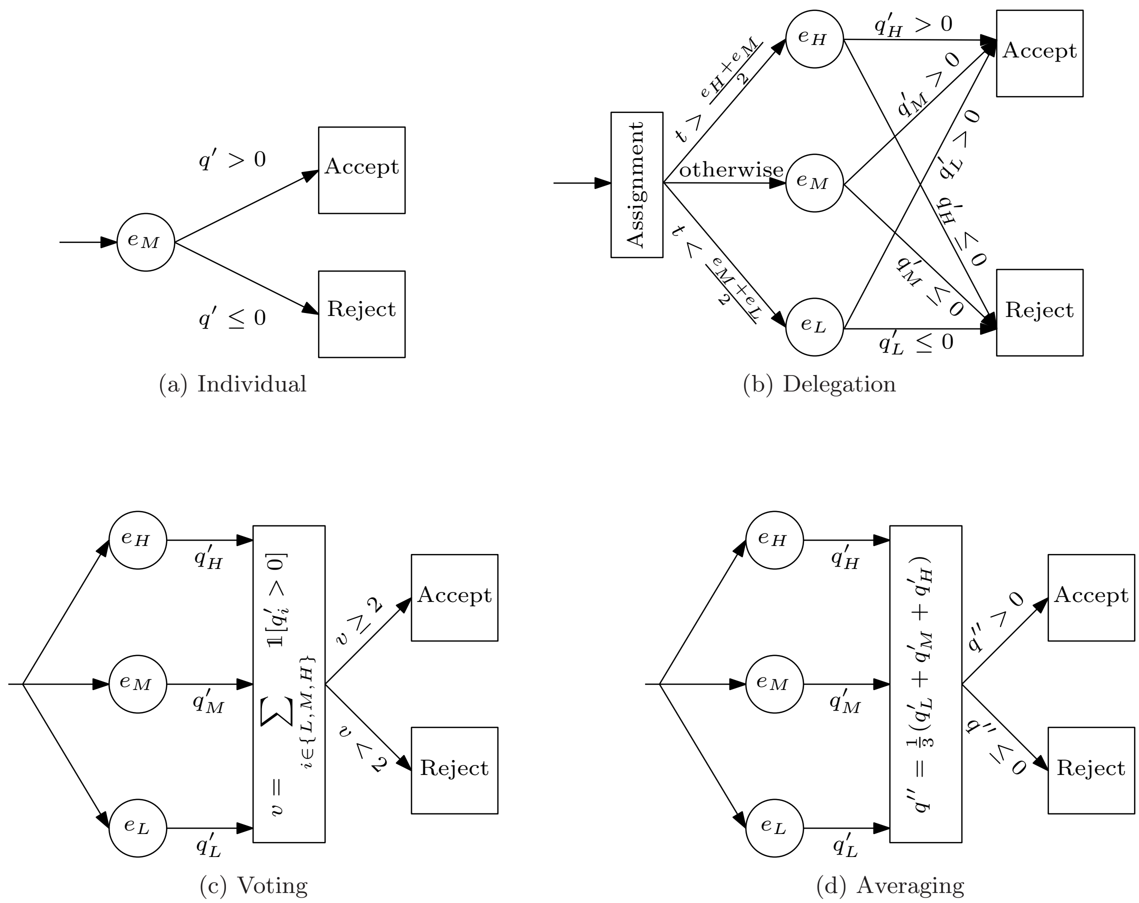


Figure 1: Graphical description of each of the structures analyzed. The input to each structure is a project (q, t).

e_M). The assigned expert then accepts the project if she perceives it as a good one and rejects it otherwise. We initially assume that the project is assigned to the manager whose expertise is closest to the project's type. Later we relax that assumption by exploring the case of imperfect delegation, as may occur when the determination of project types or individual expertise is subject to error. Delegation is the most decentralized structure, and is representative of such settings as engineering and consulting firms, where the evaluation of a project is usually assigned to the partner with the most relevant experience, or movie studios where a producer or executive might specialize in a certain movie genre and review all scripts that are pitched in that genre.

Voting (Figure 1c). Under this structure, each of the three managers evaluates the project and the organization makes a decision based on the vote of the majority (i.e., if two or more of the individuals perceive the project to have a positive quality then the project is approved, and otherwise it is rejected). This structure is representative of boards of directors, partnerships such as venture capital firms, and egalitarian top management teams.

Averaging (Figure 1d). This structure is similar to Voting; however, instead of counting votes, the structure averages the perceptions of the three individuals. If the average of the three perceived qualities is positive, then the organization accepts the project and otherwise rejects it. This structure could represent cases where committees try to fine-tune their decisions to incorporate more fully the assessments of members. For example, if two managers have a mildly negative view of a project but the third manager has an extremely positive perception of it, this structure would yield acceptance of the project. One question this research tries to answer is why this structure, which arguably takes into account more information than does Voting (continuous perceptions rather than yes/no votes) is not observed more often within organizations. It is, however, often used outside business organizations in attempts to harness the "wisdom of crowds." Examples include averaging scores from movie critics (e.g., metacritic.com) and creating a consensus earnings forecast based on estimates from multiple financial analysts (e.g., First Call).

3.4 Performance

We define organizational performance as the quality of the projects accepted by a given structure employing a given set of experts under a given external environment. We use the definitions presented previously to derive mathematically the organizational performance of each structure. Here we show how to derive this metric for the Individual structure; the metrics for the other structures can be similarly derived and are summarized in Appendix A.

Given equation (1), the probability that an individual with expertise e approves a project of quality q and type t is

$$\mathbb{P}(\text{approving a given project}) = S_{\text{ind}}(q,t;e) = 1 - \Phi(0,|t-e|)\big|_{-q}, \tag{2}$$

where $\Phi(\mu,\sigma)\big|_x$ is the cumulative distribution function of an $N(\mu,\sigma)$ evaluated at x.[8] In agreement with the previous literature (e.g., Sah and Stiglitz 1986), we call this probability the "screening function of the organization."

The previous probability is defined for a single (q,t) project. Hence, in order to obtain a performance metric for an environment of several projects it is necessary to compute expected values over the range of q-values and t-values that define that environment, i.e.,

$$\mathbb{E}(\text{quality of approved projects}) = \left(\frac{1}{\overline{t}-\underline{t}}\right)\left(\frac{1}{\overline{q}-\underline{q}}\right)\int_{\underline{t}}^{\overline{t}}\int_{\underline{q}}^{\overline{q}} qS_{\text{ind}}(q,t;e)\,\mathrm{d}q\,\mathrm{d}t. \tag{3}$$

We call this expected value the organization's *performance* in a given environment, as it corresponds to the expected accrued quality per screened project. Simply put, performance is equivalent to the total values of the accepted projects (i.e., firm profits) normalized by the number of screened projects (so that performance is comparable across organizations screening different number of projects). Because equations (2) and (3) cannot be solved analytically, we analyze the model numerically.

[8] See Csaszar (2013) for details.

4 Results

In this section we use the model to explore the relationship between organizational structure, individual expertise, and the external environment. We do this in two stages. First, we study the effect of employee expertise (i.e., the "knowledge of the organization") on the performance of each structure for a given environment. Studying firms that differ with regard to the expertise of their employees serves to elucidate the performance effects of knowledge diversity. Second, we investigate how changes in the project environment (i.e., the type of projects that constitute the stream of projects) affect the performance of each structure. Studying the effect of different project environments can shed light on how robust a structure is when faced with an unexpected change in the environment (e.g., a radical technological change).

We use graphical plots to convey the main results of the analysis in an intuitive yet precise way. Each plot shows how performance (on the y-axis) varies as a function of knowledge breadth (β, on the x-axis) for a fixed range of project types $[\underline{t}, \overline{t}]$ and project qualities $[\underline{q}, \overline{q}]$. Thus, to compare organizations within a given environment we look at one plot, and to compare across environments we look at two or more plots.

To explore the model in a way that is amenable to analysis yet representative of its behavior under a broad range of realistic conditions, we present results for a carefully chosen set of scenarios. We focus on these scenarios after extensively exploring the model and verifying that the presented results (a) capture the fundamentals of the model's behavior and (b) are qualitatively robust with respect to the exact value of the inputs (robustness checks are described in Section 4.4 and Appendix B). In other words, analyzing the model under these values is sufficient for enabling the reader to extrapolate results to other plausible scenarios.

In general, two regularities allow us to keep the number of scenarios presented to a minimum. First, the effect of varying the range of project qualities $[\underline{q}, \overline{q}]$ is straightforward: as the quality of the portfolio increases, the performance of all the structures increases monotonically and without affecting the performance ranking of the structures. In the limit (when project qualities become extremely positive or extremely negative), all structures perform identically (accepting or rejecting all projects, respectively). Thus, to understand the behavior of the model, it is enough to study it under one range of qualities (in the ensuing analyses we use $q \sim U[-5, 5]$). The second regularity is that, because the noise in individual perception is a function of the difference between project type and individual expertise (by equation (1), which sets noise proportional to $|t - e|$), to understand the behavior of the model it is enough to vary one of the elements of the difference (t or e) while keeping the other fixed; in most of our analyses, we vary the expertise while keeping the range of project types fixed.

Within each plotted scenario (defined by a range of project types t and qualities q), we explore the behavior of the different structures while varying the degree of expertise heterogeneity within each structure. As mentioned previously, in most of our analyses the respective expertises of the three managers employed by Delegation, Voting, and Averaging are parameterized by the expertise e_M of the middle manager and knowledge breadth β such that $e_L = e_M - \beta$ and $e_H = e_M + \beta$.

In the first scenario analyzed, we set the expertise of the middle manager at exactly the center of the range of project types (i.e., $e_M = \frac{\bar{t}-\underline{t}}{2}$). This choice of e_M minimizes the expected perception errors of that manager. This assumption is relaxed later, but we consider this choice of e_M a reasonable starting point because mechanisms such as competition or learning could allow firms to discover that this is an optimal position for e_M. We vary knowledge breadth (β) from zero (i.e., all individuals are identical) to the value at which the three experts are maximally different but still remain within the range of the project types (e.g., if $t \sim U[0, 10]$ and $e_M = 5$, then $\beta_{\max} = 5$ since this value leads to $e_L = 0$ and $e_H = 10$).

We structure the analysis as follows. First, we familiarize the reader with the inner workings of each structure by focusing on a base case. This first step helps uncover the basic mechanisms that relate knowledge breadth (β) to performance for each structure. Second, we explore the effect of shifting the range of project types while keeping the expertise of the individuals fixed. Third, we explore what happens to the performance of Delegation when we relax the assumption of perfect assignment of projects. Finally, we describe the robustness of the findings and important model extensions. More general implications of the model are discussed in Section 5.

4.1 Base Case

Figure 2 shows the performance of the four structures as a function of knowledge breadth β in a single environment ($q \sim U[-5, 5]$ and $t \sim U[0, 10]$) and under the assumption that the middle manager is located at the middle of the project types being considered ($e_M = 5$). To get a sense about the of the values of performance in the y-axis, note that a random decision maker would get a performance of zero, while an omniscient one (someone approving all the good projects and rejecting all the bad ones) would get a performance of 1.25.[9] A first look at this figure reveals several nontrivial relationships, including noticeable differences between Averaging and Voting in addition to the strong performance of Delegation as knowledge breadth increases. The following paragraphs delve into the mechanisms explaining these and other differences in performance across structures.

A baseline observation from Figure 2, which also holds for all subsequent figures, is that the performance of the Individual structure is flat with respect to knowledge breadth. This follows directly from the definition of this structure, in which β does not play any role—this structure involves only one individual, whose expertise is e_M. Because this structure is so simple (in terms of both structure and performance) and because it serves as a building block of all the other structures, it makes a natural benchmark against which to compare the performance of those other structures. This figure shows that the Individual structure is the poorest performer (with one exception in the lower right of the figure, which we discuss later). Thus, in our model, going from organizations of one individual to organizations of three individuals is almost always associated with a positive effect on performance.

Averaging versus Voting. Perhaps the most striking observation from Figure 2 is that, although

[9] From equation (3), the performance of the omniscient decision maker is $\frac{1}{10-0} \frac{1}{5-(-5)} \int_0^{10} \int_{-5}^{5} q \mathbb{1}[q > 0] \, dq \, dt = 1.25$.

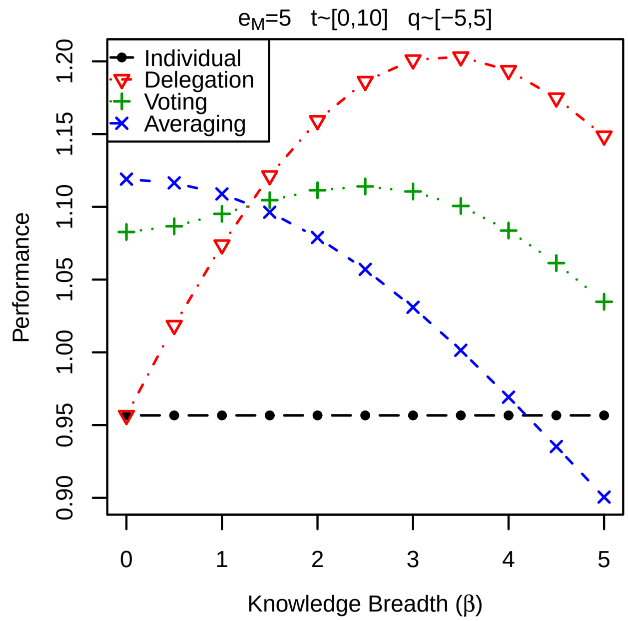


Figure 2: Performance of the four structures in the base case.

Voting and Averaging are intuitively similar, the relationship between performance and knowledge breadth in these two structures is quite different. The performance of Averaging peaks when its managers are the most similar (i.e., with no knowledge breadth or $\beta = 0$), and it decreases rapidly as knowledge breadth increases. In contrast, Voting is less sensitive to knowledge breadth and exhibits a nonmonotonic behavior: performance peaks at a moderate level of β before declining slightly. The performance difference between Voting and Averaging is particularly interesting in that everybody we informally surveyed beforehand believed that any difference between the performance of the two structures would be to the advantage of Averaging, which takes more information into account (Averaging combines individual perceptions whereas Voting combines votes, which are nothing more than discretized versions of those individual perceptions). Moreover, this pro-Averaging argument seemed consistent with the well-known concept of applied statistics that running a statistical procedure (e.g., a regression) on continuous variables, rather than on binary versions of them, produces better estimates. But these a priori arguments proved to be deceptive.

The logic explaining the counterintuitive difference between Voting and Averaging is that Averaging overweights perceptions of the less suitably informed individuals. For example, if experts $e_L = 1$, $e_M = 5$, and $e_H = 9$ review a project of type $t = 7$ and quality $q = 1$, then experts e_M and e_H would likely perceive the project with the least noise, putting their estimates very close to $q = 1$. Meanwhile, expert e_L would be as likely to have a q' of -5 as $+7$ (i.e., $q \pm |t-e| = 1 \pm 6$), making the decision to accept or reject the project highly dependent on the specific opinion of e_L. Thus, under Averaging, the perception of the least-suited individual can throw the organization-level decision completely off balance. In contrast, Voting has the property that, no matter how mistaken is the perception of any individual, her effect on the final decision is capped: she can only affect the vote count by one vote.

The only case where Averaging surpasses Voting is when knowledge breadth is close to zero (i.e., $\beta < 1$ in Figure 2). In this case, all the individuals are similarly qualified to evaluate the project, so no individual has undue power to throw the averaging process off balance. When individuals are identical and thus unequal noise ceases to be an issue, the additional information carried in the continuous signals of Averaging (versus the discrete signals of Voting) is beneficial. In other words, when β is low, the continuous signals used by Averaging do not come at a significant cost and the informational advantage of these signals becomes observable.[10] An intriguing implication of the possibility that Averaging may underperform Voting is that groups should not necessarily use all the information at their disposal. Sometimes, less information is better (i.e., discrete votes instead of full-fledged perceptions).

An application of the previous comparison is that if the organization designer is unsure about the actual expertise of the organization's members then he may be better-off choosing Voting rather than Averaging. In the opposite case, an organization designer who can perfectly choose the expertise of its members may be better-off employing members of identical expertise and using Averaging rather than Voting. An instance of this advantageous use of Averaging could be in Olympic competitions, where scores are usually computed by averaging the scores of the judges (rather than by counting votes); as arguably, the organizers of the games can choose many judges that have similar expertise (e.g., ex-Olympians). In most other settings—conspicuously in business and government—where perfect calibration to a well-defined task is unlikely to occur or hard to assess, the organization designer might better employ Voting over Averaging. In sum, Voting structures are robust to miscalibrated members. This characteristic may explain why voting is so prevalent in settings that face a variety of environments and high member turnover, such as boards of directors and legislative bodies.

Delegation. We now explore the behavior of Delegation, whose performance (as shown in Figure 2) also exhibits a distinct path—in this case, an inverted-U shape. We describe the performance of Delegation in three steps. Initially, under knowledge homogeneity ($\beta = 0$), Delegation performs identically to the Individual structure because any individual receiving a delegated project is identical to the only member of the Individual structure. Then, as β increases, so does the performance of Delegation. As the decision makers become spread over a larger portion of the knowledge range, it becomes more likely that the expertise of the manager receiving the delegated project is close to the type of the project. Finally, the relationship between performance and β peaks at an intermediate level of β. To understand why the maximal performance of Delegation occurs there, imagine three goalkeepers jointly trying to cover a goal that is 6 meters wide. Simple arithmetic shows that, if the goalkeepers position themselves 0, 3, and 6 meters from the left post then, in the worst-case scenario, a goalkeeper will be 1.5 meters from the ball when it crosses the goal line. However, if the goalkeepers locate at positions 1, 3, and 5, then one of them will be at most 1 meter from the

[10] In statistical terms, continuous signals have an informational advantage over discrete signals because they allow for errors to be cancelled out more efficiently or, equivalently, because they have a greater asymptotic accuracy (Casella and Berger 2002:470).

ball. A similar reasoning explains why, under Delegation, the optimal knowledge breadth is less than the maximal β (i.e., the value of β that would locate experts at the corners of the range of project types).[11]

In Figure 2 Delegation becomes the best performing structure at around $\beta = 1.3$. One may wonder if this level of knowledge heterogeneity is not so high that, in reality, no one would plausibly consider using egalitarian rules such as averaging or voting after that level of β (because as β increases, the least suitable individual becomes patently farther and farther off from the project being evaluated). This plausibility condition can be formally analyzed by comparing the accuracy of the best and worst expert at the point where Delegation becomes the superior structure. Following Soll and Larrick (2009), we measure this relative accuracy by computing the mean absolute error (MAE) of the worst expert versus the MAE of the best expert. At the crossover point in Figure 2, the MAE of the least suitable expert is about 1.6 times larger than of the MAE of the best expert, which seems a plausible ratio to encounter in a real organization.[12]

Comparing structures. An interesting perspective emerges from reading Figure 2 normatively: consistent with the main tenet of contingency theory (Lawrence and Lorsch 1967), there is no single best organizational structure. In fact, Delegation, Voting, and Averaging each can be the best performer depending on the distribution of knowledge breadth (β). With perfect information about the knowledge breadth of the organization, one should Average (when β is low) or Delegate (otherwise). Thus, if β is known, Voting is dominated by the combination of Averaging and Delegation. But with imperfect information about the knowledge breadth of the organization one might prefer to use Voting as it has lower variance, and in some regions a higher mean than randomly choosing between Averaging and Delegation.[13] In other words, Averaging performs best when β is low, Delegation performs best when β is high, and Voting may be the best choice when β is uncertain (which could happen if the organization designer is uncertain about the individuals' expertises or the projects' types). We further discuss across-structure comparisons in Section 5.1.

4.2 Changing Environment

Here we explore how these structures perform when the environment changes. Specifically, in this section we explore the case of a shift in the type of projects reviewed by the firm while the expertise of its members remains fixed. This type of change is akin to the effect of a radical new technology, such as the shift from analog to digital photography (Tripsas and Gavetti 2000). For example, at Polaroid during this transformation, expertise related to hiring software engineers became more important while expertise related to hiring chemists became less important. At the same time, because of organizational inertia and the unexpected nature of the change, an incumbent firm

[11] The performance of Delegation is highly contingent on the ability to assign the project to the proper expert, an issue we explore in Section 4.3.

[12] For projects uniformly distributed in the range $[\underline{t}, \overline{t}] \times [\underline{q}, \overline{q}]$ and structures with three experts, the ratio of the MAEs corresponds to $\frac{\mathrm{MAE}_{\mathrm{Worst}}}{\mathrm{MAE}_{\mathrm{Best}}} = \frac{\int_{\underline{t}}^{\overline{t}} \max\{|t-(e_M-\beta)|, |t-e_M|, |t-(e_M+\beta)|\}\,\mathrm{d}t}{\int_{\underline{t}}^{\overline{t}} \min\{|t-(e_M-\beta)|, |t-e_M|, |t-(e_M+\beta)|\}\,\mathrm{d}t}$.

[13] We thank an anonymous reviewer for highlighting this last point.

like Polaroid could not adapt its expertise instantly and so remained stuck with its previous set of experts, who had been selected based on the older, analog environment. The analyses in this section shed light on what structures and levels of knowledge breadth are better able to cope with radical environmental changes. In other words, we now study how structure and knowledge breadth can protect firms that cannot instantly change their managers every time the environment varies.

The panels in Figure 3 show four snapshots of performance as the environment gradually varies from the base case ($t \sim U[0, 10]$ in the first panel) to a radically different environment ($t \sim U[15, 25]$ in last panel). A first observation from this figure is that the overall performance of the structures (i.e., the ranges on the y-axes) decreases as we move from panel (a) to panel (d). This happens because, as we advance through the panels, the project types begin to drift away from the expertise of the firm, which remains fixed (at $e_L = 5 - \beta$, $e_M = 5$, and $e_H = 5 + \beta$). Thus, for example, in the last panel, experts centered around $e_M = 5$ must screen projects with types between 15 and 25, which leads to large evaluation errors and a consequent effect on performance.

Another observation from Figure 3 is that, at each step, Averaging increasingly trumps Delegation as the best performer at low and middle levels of expertise breadth (i.e., in the first panel Averaging dominates until $\beta = 1$, whereas in the last panel Averaging dominates over the entire range of β). The rationale is that, as the environment shifts farther and farther away from the decision makers within the firm, the errors made by the decision makers in evaluating projects become increasingly similar in their distributions (when the shift is extreme, the three decision makers become similarly incompetent in their ability to assess the projects under consideration). This convergence improves the relative performance of Averaging over Delegation, as averaging similarly uninformed opinions cancels out some of the error, whereas delegating to a single uninformed individual does not. Hence, a group could make comparatively good decisions even if none of its members has the appropriate expertise to make a good decision individually.

The explanation for the effect of environmental change on Delegation is quite different. Although Delegation loses some terrain against Averaging (and Voting) in the first three panels of Figure 3, Delegation remains the top performer when knowledge breadth (β) is high. This is because, when β is high, the expert who receives the delegated project is better equipped to deal with the new project than is the uneven mix of experts who participate in Averaging or Voting. To make the argument clearer, suppose that experts $e_L = 0$, $e_M = 5$, and $e_H = 10$ are evaluating a project of type $t = 12$, which is slightly above the level of e_H's expertise. Delegation would send the project to e_H, whose noisy perception would have a standard deviation of 2 ($= |10 - 12|$). Averaging or Voting would additionally use the opinions of the two other experts, whose noisy signals have a much higher standard deviation (of $|0 - 12| = 12$ and $|5 - 12| = 7$) and would thus throw off balance the relatively accurate perception of the other individual rather than cancel out its error. Thus, Delegation performs quite well in a changing environment as long as one of the decision makers possesses expertise that is within or near the range of project types under consideration. As the expertise of the closest decision maker (e_H) becomes less helpful in the new environment, the performance of Delegation falls below that of Averaging.

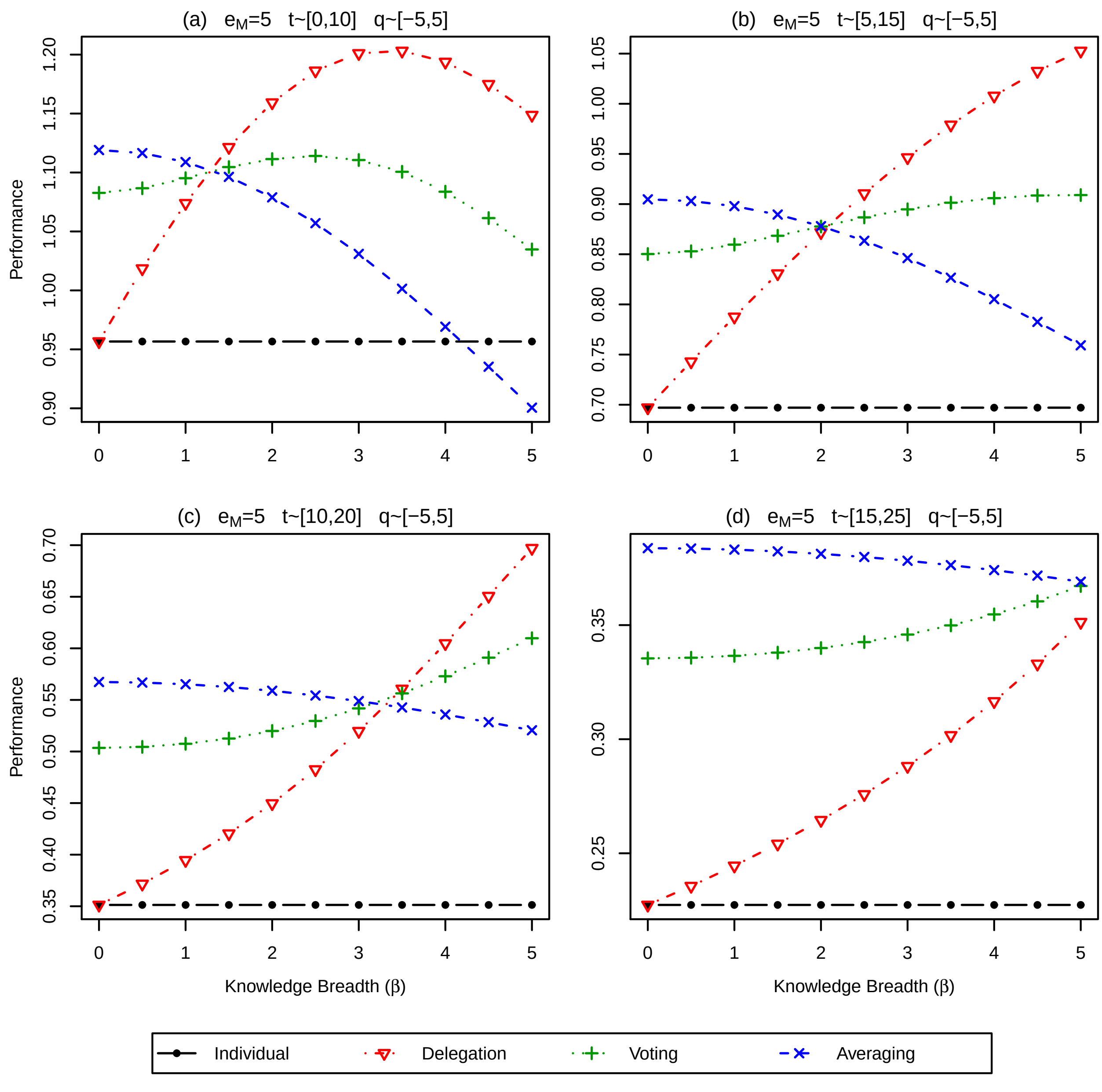


Figure 3: Four snapshots of performance as the range of project types changes (from $t \sim U[0, 10]$ in the panel (a) to $t \sim U[15, 25]$ in the panel (d)).

One implication of Figure 3 is that Delegation combined with high breadth of expertise is an organization design that is robust to unexpected shifts in the range of project types—except at the most extreme levels of change (i.e., Figure 3d). Observe that under the extreme change scenario, none of the firm's original experts (e between 0 and 10) is of much use in the new space (t between 15 and 25). However, a scenario like this is only descriptive of the most extreme environmental changes. Most changes in the real world probably do not reach these levels. For instance, even after the shift to digital photography, Polaroid's competencies in camera design and lens technology were still relevant. Therefore, one empirical implication of our model is that Delegation (when knowledge breadth is high) exhibits comparatively higher performance during most periods of environmental change.

4.3 Imperfect Delegation

Delegation has an important difference with respect to the other structures: it incorporates a first stage during which the project is assigned to a single decision maker. So far, the model has assumed that projects are perfectly assigned to the expert whose expertise is closest to the type of the received project. But in real life this assignment process can be imperfect. Errors in assignment can occur, for example, if project type or individual expertise are hard to assess, if there is not a reliable process for matching projects to experts, if there are political reasons for not assigning projects to the right experts, or if the right expert is unavailable. Studying the sensitivity of Delegation to errors of assignment is particularly relevant because our analyses so far have identified a wide range of cases where Delegation is the top-performing structure.

To account for possible assignment imperfections in Delegation, we introduce a new parameter, r, which captures error in the assignment process. When $r = 0$, projects are always assigned to the best expert, while when $r = 1$ projects are randomly assigned among the three experts. Intermediate values of r interpolate between these two extremes. More formally, parameter r affects the delegation process as follows: with probability $\frac{r}{3}$ the project is erroneously assigned to any of the two least suitable experts and otherwise (with probability $1 - \frac{2r}{3}$) the incoming project is assigned to the best-suited expert. A mathematical definition of imperfect delegation appears in Appendix A.

Figure 4 plots the performance of Delegation for different values of r under the base case environment ($t \sim U[0, 10]$).[14] The main observation from this figure is that Delegation is quite sensitive to assignment errors. For example, if the error rate rises from 0 to 20% (compare lines $r = 0$ and $r = 0.2$ in Figure 4) then the peak performance of Delegation falls from roughly 1.20 to 1.12. Comparing Figures 2 and 4 shows that when r is slightly above 0.2, Delegation falls below the peaks for both Averaging (for low β) and Voting (for high and medium β). An error rate slightly above 0.6 would render Delegation the worst-performing of the structures (i.e., even worse than the Individual structure for high levels of β). Delegation's sensitivity to assignment errors is an important characteristic to keep in mind when considering managerial implications.

[14] The results for other environments are qualitatively equivalent.

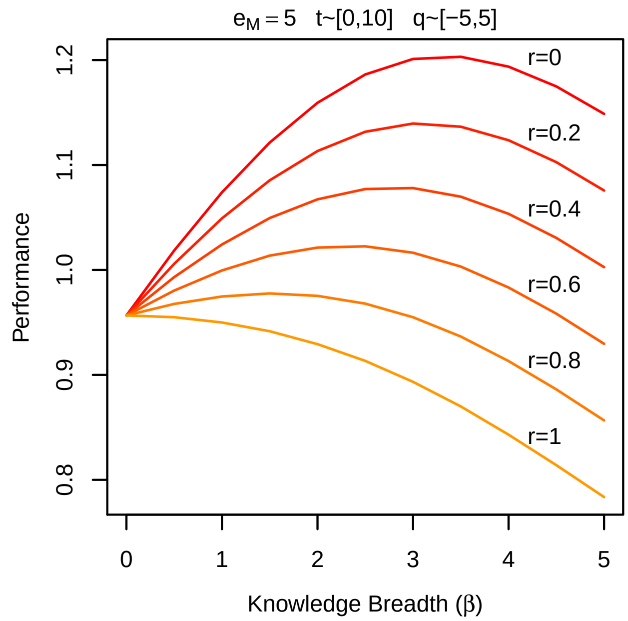


Figure 4: Performance of Delegation under imperfect assignment, where parameter r denotes the error rate.

4.4 Robustness Checks and Extensions

We studied the robustness of our findings with respect to a broad range of parameter values and model specifications. These tests included: (a) setting the expertise of individuals in a random (rather than a symmetric) fashion, (b) adding a base error rate to the individual perceptions, (c) scaling the individuals' errors, and (d) changing the probability distributions of the project types. In general, our results are qualitatively robust to changes in all of these assumptions. A representative set of the robustness tests is shown in Appendix B.

We also studied three extensions to our model. These extensions provide some interesting results and also illustrate how our model could be used as a platform to explore additional issues of information aggregation in organizations. We include these analyses as "extensions" and not as part of the main results of the paper, as each of them could potentially be analyzed in more detail, but such detailed analysis is outside the scope of this initial paper. Nevertheless, we believe that there is value in briefly describing each of these three modeling extensions and the main results that emerge from them. Figures and additional details are given in Appendix C.

The first extension we consider is the effect of correlated errors among the experts, an analysis in the spirit of Clemen and Winkler (1985), which is representative of phenomena such as groupthink (Janis 1972) or herding (Bikhchandani et al. 1992). The results of this analysis (see Figure C.1 in Appendix C) show that the performance of both Averaging and Voting declines rapidly as the errors of the decision makers become more correlated. This suggests that error correlation among experts can harm the performance of Voting and Averaging, in a similar way that assignment errors can harm the performance of Delegation.

The second extension we consider is to place optimal weights on the experts' opinions in Voting and Averaging, an analysis in the spirit of Ben-Yashar and Nitzan (1997). Calculating these optimal weights is computationally complex and requires precise information on the incoming project and individuals' expertises. It is not therefore particularly plausible from a behavioral point of view, but comparing the performance of the simple structures in this paper against optimal benchmarks is informative. The results of this analysis (see Figure C.2 in Appendix C) show that optimal Averaging dominates all the other structures for all levels of β. This is reasonable, as Delegation is a particular case of optimal Averaging (i.e., put all the weight in the best expert), and in Section 4.1 we had shown that the combination of Averaging and Delegation dominated performance in the base case. More interestingly, the performance of optimal Averaging is never more than 5% above the performance of the best among the simple versions of Averaging and Delegation studied throughout the paper. This is in line with classic results on the power of simple heuristics (e.g., Einhorn and Hogarth 1975, Dawes 1979, Gigerenzer and Goldstein 1996).

The last extension considers individuals who do not report their true opinions, but an adjusted version of them. This adjustment process captures the idea that an individual may realize that she is not competent to analyze a given project and thus may decide to "correct" her estimate by reporting an estimate that is closer to the typical project that the organization receives. We model the correction phase by updating the individual's perception q' into a convex combination between the perception and the mean project, that it, $q'_{\text{corrected}} = f(\cdot)q' + (1 - f(\cdot))\frac{\underline{q}+\overline{q}}{2}$. We interpret the function $f(\cdot)$ as self-confidence (i.e., when $f(\cdot) = 1$ the individual only takes into account her own perception), akin to the parameter weight on self, ws, in Soll and Larrick (2009). We consider two shapes for the function $f(\cdot)$: (a) one that decreases monotonically as the expert is farther away from the type of the incoming project (i.e., $f(\cdot)$ decreases with $|t - e|$), and (b) one that is U-shaped with respect to $|t-e|$ (e.g., a competent expert reports what she perceives, a half-competent expert reports an average between what she perceives and the average project, and an incompetent expert does not know she is incompetent and simply reports what she perceives). Alternative (a) is consistent with Bayesian decision making, and alternative (b) is consistent with empirical research on perceptions of competence (Kruger and Dunning 1999).

The results for this extension (see Figure C.3 in Appendix C) show that having individuals that update their estimates affects Averaging the most. Under alternative (a), the performance of Averaging improves, as the potentially damaging opinions of the least-informed individuals are now made less extreme. Under alternative (b), the performance of Averaging falls drastically when knowledge breadth (β) is high. This is reasonable, as the uninformed individuals are unaware of their incompetence and their unfiltered opinions throw Averaging off balance. Under both alternatives, the performance of Voting remains robust, and Delegation remains the best performer at higher levels of β. In sum, how individuals behave when they are incompetent affects structures differently. Voting remains robust, while the relative performance of Averaging can vary wildly.

5 Discussion

This study introduces a mathematical model involving organizational structure, individual expertise, and the external environment. The model extends previous models of information aggregation by infusing organizational elements such as heterogeneous knowledge and a changing external environment. We use the model to compare the performance of four decision-making structures: an Individual decision maker, a Delegation process, a Voting body, and an Averaging body. The results demonstrate that structure, expertise, and environment are tightly interconnected in their effect on organizational performance. Below we discuss practical uses of the model, as well as connections between our work and crowd-based decision making and organizational adaptation.

5.1 Managerial Application

An application of the model consists in recommending a decision-making structure based on given values for the parameters. A concise way to organize model recommendations is by using decision trees. Figure 5 shows one such decision tree, which encapsulates the structure comparison for the base case (Section 4.1) and recommends the most appropriate structure—assuming that the environment is static and that knowledge breadth is given (i.e., managers of the organization cannot be replaced or retrained). The decision tree takes into account four conditions (depicted as boxes), which we describe briefly, starting from the top of the decision tree. The first condition is a cost–benefit one: if the cost of employing three decision makers does not compensate for its benefits, then the Individual structure is preferable (this argument assumes that three decision makers are costlier than one). The second and third conditions relate to knowledge breadth (β): if β is unknown (as when expertise is difficult or costly to assess), then Voting offers a robust performance; if β is low, then Averaging offers the best performance; and if β is high, then which structure is preferable depends on the likelihood of delegation errors. If delegation errors are likely, then Voting is a safe bet; otherwise, Delegation should be used. The figure shows that each of the four structures can be a viable choice under the appropriate set of conditions (i.e., the four structures appear as terminal nodes of the tree), which underscores the contingent nature of organizational design (Lawrence and Lorsch 1967, Nickerson and Zenger 2004).

Similar decision trees can be constructed for the three remaining combinations that stem from considering knowledge breadth as given or controllable and the environment as stable or changing as well as for the extensions mentioned in Section 4.4. We do not present these trees here for the sake of brevity (in any case, they can be inferred from the Results section and Appendix C).

The construction of the model implies some important boundary conditions on the results. First, in the spirit of Hotelling (1929), we assume that it is possible to measure a unidimensional distance between project type and individual expertise. Theoretically (as explained in footnote 5) this assumption is not very restrictive. But in practice, it could be that the distance between e and t is hard to measure accurately, leading to noisy assignments as described in the context of imperfect Delegation (Section 4.3). Second, our model assumes no interactions between projects, in terms of

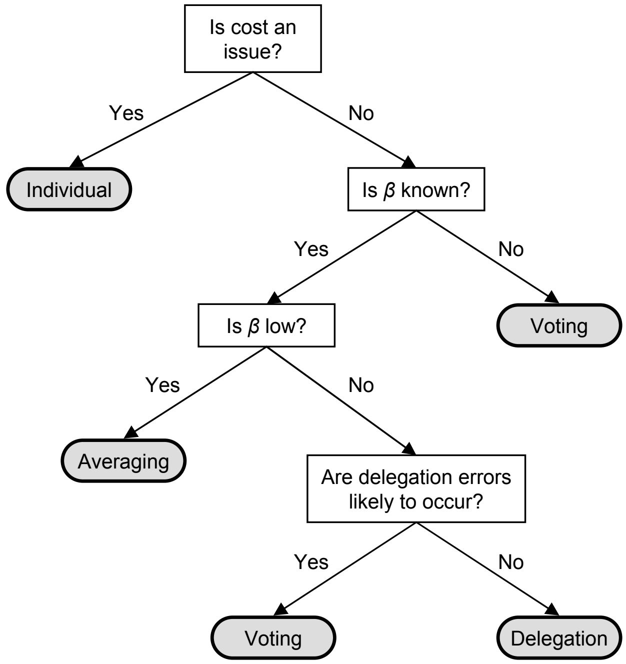


Figure 5: Decision tree that recommends the best structure as a function of knowledge breadth (β) and delegation errors (r), and assuming a static environment.

capacity constraints, tradeoffs, or complementarities. The presence of significant interactions might increase the appeal of centralized structures (Siggelkow and Rivkin 2005). Third, our model does not incorporate the cost of decision makers. Cost considerations could affect the overall level of the performance curves, and thus make some structures unprofitable. Fourth, we have focused here on the process of alternative evaluation only; thus the model has nothing to say about how structure affects the performance of other stages of organizational decision making.

5.2 Delegation or "The Wisdom of Crowds"?

The "wisdom of crowds" (Surowiecki 2004, Kameda et al. 2011) and related concepts such as prediction markets (Wolfers and Zitzewitz 2004) have received significant coverage in the practitioner literature. The general tone of this coverage is that choices made by crowds are superior to the results of traditional, delegative decision making in firms. For example, a recent article in the *Wall Street Journal* (Murray 2010) postulates that crowd decision making would cause "the end of management." Yet, casual observation indicates that delegation is probably the most prevalent form of decision making in firms. Are firms wrong to choose delegation over crowds?

The model in this paper can shed light on when delegation is superior to crowd-based decision making and vice versa. To compare Delegation with a crowd, we created Voting and Averaging models with increasingly larger numbers of decision-makers to determine at what point these models dominate Delegation. Figure 6 shows that, in the base case scenario, Voting with nine individuals dominates Delegation, while Averaging requires fifteen individuals. The numbers in this analysis should not be read literally, as they depend on a number of assumptions (i.e., we used the base case assumptions, located the N experts evenly over the knowledge range, and assumed uncorrelated errors). In real life, although the number of individuals in a crowd may be much larger than nine or fifteen, their knowledge breadth may be narrow or their errors correlated (as discussed in Section 4.4 and in Appendix C.1), thus acting equivalently to a small group and greatly reducing the overall performance of the crowd. Also, in many situations accessing the crowd may be impractical for cost or confidentiality reasons. But the main message of this analysis is suggestive: a well-chosen crowd can surpass the performance of delegation. Thus, this model adds to the conversation (e.g., Hastie and Kameda 2005, Kameda et al. 2011) about distinguishing between the wisdom (Surowiecki 2004) and the madness (Mackay 1852) of crowds.

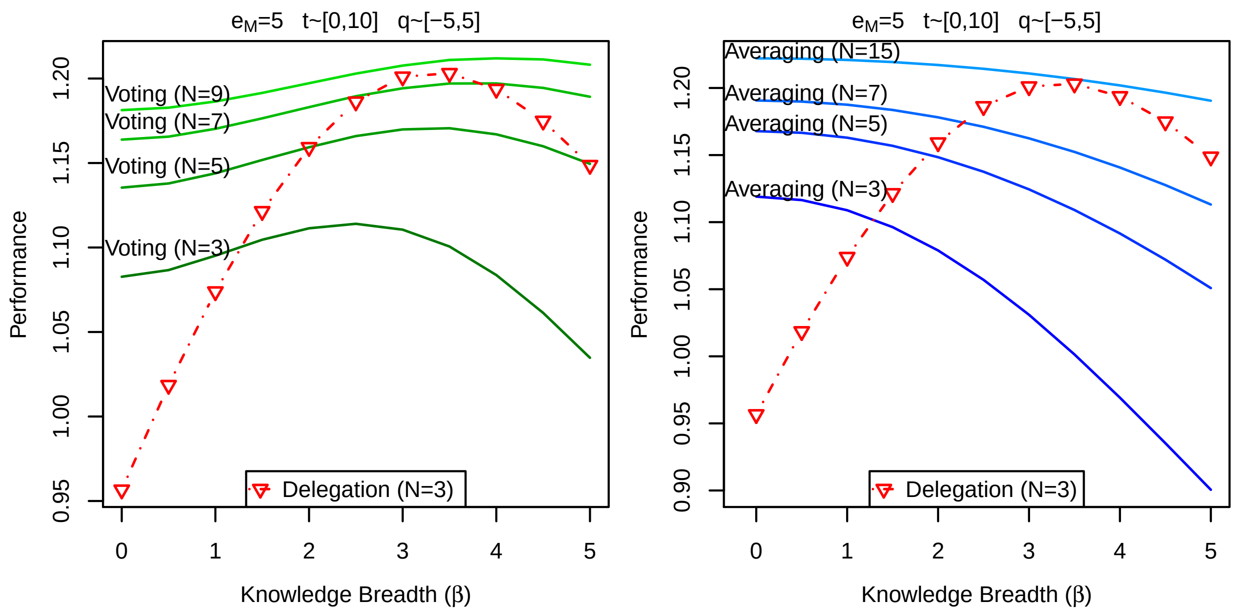


Figure 6: Performance of Voting and Averaging as a function of the number of the number of decision makers (N).

Another way of thinking about crowds versus delegation is in terms of "crowdlike" decision-making within the firm. Averaging and Voting are more crowdlike than Delegation, as they take into account the opinion of more individuals and they aggregate this information in a more egalitarian way. In our models with three decision-makers, delegation outperforms crowdlike structures (Voting or Averaging) when the projects assessed by the firm are well matched by the expertise of the delegated individual (in terms of the model, when β is moderate or high), when significant environmental shocks are unlikely to make existing firm expertise obsolete, and when assignment errors are relatively rare (r is low). Otherwise, one of the crowdlike structures is preferable. We conjecture that the conditions under which Delegation outperforms the other structures have been prevalent throughout business history, so that (in the sense of organizational evolution) Delegation has become the default decision structure for firms. A further managerial implication is that crowdlike mechanisms should be considered in settings that deviate from the conditions above.

A related question is why, when firms do use a crowdlike mechanism, it is usually by means of Voting and only rarely by means of Averaging. Prior work has suggested that the Averaging structure can be costly: the process of arriving at a specific number can be more time-consuming and cognitively demanding than Voting's "thumbs up–thumbs down" approach or Delegation's division of labor (Hastie and Kameda 2005). Similarly, the value of some potential projects may be difficult to distill down to a single value, making Averaging impractical. Our model suggests two additional reasons why most firms make little use of Averaging structures. First, it may be challenging to ensure that all decision makers are similar in terms of expertise (i.e., β is close to 0 in Figure 2), either because doing so is costly or because organizations (and their members) will change over time. Second, Averaging is more susceptible than Voting to agency concerns, since Averaging would overweight an outlier opinion even if that opinion were based on personal preferences and not on facts. In contrast, Voting is less susceptible to this bias because the negative effect of a dishonest individual is capped at one vote.

5.3 Structure and Environmental Change

A long-standing question in management is how firms can use organizational structure to cope with environmental change. Some of the early structures that have been suggested in this regard are Organic (Woodward 1965), Prospector (Miles and Snow 1978), and Adhocracy (Mintzberg 1979). Our work here suggests that there is no single correct structure to deal with environmental change and that the proper choice depends on internal and external contingencies—namely, the knowledge possessed by the organization's members and the knowledge required by the environment. Specifically, for a firm possessing knowledge that remains relevant in the changing environment, Delegation makes the most sense. This finding aligns with work on technological change, which suggests that firms could adapt more successfully by placing authority in the hands of, for example, scientists with directly relevant knowledge instead of executives who may be wedded to existing business models (Christensen and Bower 1996, Tripsas and Gavetti 2000, Eggers and Kaplan 2009). However, if the firm does not possess the knowledge relevant to the new environment (either because the individuals' expertise is extremely homogeneous or because the environmental shift is radical), then it would be more advantageous for the firm to employ Voting or Averaging.

5.4 Future Work

The research presented here can be extended in several directions. Possible avenues for *empirical* work include: (a) testing the model's predictions (e.g., test the hypotheses implicit in the decision tree of Figure 5); (b) developing ethnographic studies to better understand how real organizations use and misuse the mechanisms proposed here; and (c) measuring the effect sizes associated with the model's parameters, as well as the variance explained by information aggregation versus other firm- and industry-level characteristics.

Possible avenues for *theoretical* work include: (i) predicting other performance metrics, such as measures of variance or risk, and errors of omission and commission; (ii) incorporating other processes into the model, such as incentives and learning; and (iii) comparing the performance of a broader set of structures, such as organizations with dynamic decision rules or complex network structures. We believe that models of information aggregation, such as the one presented in the current study, hold great promise for the study of organizations because they are fully aligned with the main tenets of behavioral theory of the firm: that an organization's main task is to process information and that organizational structure determines how information is aggregated.

5.5 Conclusion

Our study constitutes a step toward better understanding the links between information aggregation and the process of organizational decision making by uncovering complex interdependencies between organizational structure, individual expertise, and environmental change. From a practical standpoint, our results shed light on which structure to use when. Some of the questions we have addressed are: why delegation is a common structure for organizations in more stable environments while voting is more common under changing environments or memberships; how structure can be used to cope with radical environmental change; and when organizational decision making is better-off using less (rather than more) information from its members. From a theoretical standpoint, we address several calls to enrich our understanding of organizations by describing a causal mechanism that links the structure of organizational decision making to organizational performance. Overall, our results underscore the critical role of information aggregation on organizations.

References

Aristotle. c.330 BCE/1984. Politics. J. Barnes, ed., *Complete Works of Aristotle: The Revised Oxford Translation, vol. 2*. Princeton University Press, Princeton, NJ.

Arrow, K. J. 1951. *Social Choice and Individual Values*. Yale University Press, New Haven, CT.

Ben-Yashar, R. C., S. I. Nitzan. 1997. The optimal decision rule for fixed-size committees in dichotomous choice situations: The general result. *International Economic Review* **38**(1) 175–186.

Bikhchandani, S., D. Hirshleifer, I. Welch. 1992. A theory of fads, fashion, custom, and cultural-change as informational cascades. *Journal of Political Economy* **100**(5) 992–1026.

Bower, J. L. 1970. *Managing the Resource Allocation Process: A Study of Corporate Planning and Investment*. Division of Research, Graduate School of Business Administration, Harvard University, Boston, MA.

Brunswik, E. 1952. *The Conceptual Framework of Psychology*. University of Chicago Press, Chicago, IL.

Burton, R. M., B. Obel. 2004. *Strategic Organizational Diagnosis and Design: The Dynamics of Fit*. 3rd ed. Kluwer, Boston, MA.

Carley, K. M., Z. A. Lin. 1997. A theoretical study of organizational performance under information distortion. *Management Science* **43**(7) 976–997.

Casella, G., R. L. Berger. 2002. *Statistical Inference*. 2nd ed. Duxbury, Pacific Grove, CA.

Christensen, C. M., J. L. Bower. 1996. Customer power, strategic investment, and the failure of leading firms. *Strategic Management Journal* **17**(3) 197–218.

Christensen, M., T. Knudsen. 2010. Design of decision making-organizations. *Management Science* **56**(1) 71–89.

Clemen, R. T., R. L. Winkler. 1985. Limits for the precision and value of information from dependent sources. *Operations Research* **33**(2) 427–442.

Cohen, M. D., J. G. March, J. P. Olsen. 1972. A garbage can model of organizational choice. *Administrative Science Quarterly* **17** 1–25.

Condorcet, J. A. N. 1785/1994. *Foundations of Social Choice and Political Theory*. E. Elgar, Brookfield, VT. Translated and edited by I. McLean and F. Hewitt.

Csaszar, F. A. 2012. Organizational structure as a determinant of performance: Evidence from mutual funds. *Strategic Management Journal* **33**(6) 611–632.

Csaszar, F. A. 2013. An efficient frontier in organization design: Organizational structure as a determinant of exploration and exploitation. *Organization Science* Forthcoming.

Cyert, R. M., J. G. March. 1963. *A Behavioral Theory of the Firm*. Prentice-Hall, Englewood Cliffs, NJ.

Dawes, R. M. 1979. The robust beauty of improper linear-models in decision-making. *American Psychologist* **34**(7) 571–582.

Eggers, J. P., S. Kaplan. 2009. Cognition and renewal: Comparing CEO and organizational effects on incumbent adaptation to technical change. *Organization Science* **20**(2) 461–477.

Einhorn, H. J., R. M. Hogarth. 1975. Unit weighting schemes for decision making. *Organizational Behavior and Human Performance* **13**(2) 171–192.

Einhorn, H. J., R. M. Hogarth, E. Klempner. 1977. Quality of group judgment. *Psychological Bulletin* **84**(1) 158–172.

Fang, C., J. Lee, M. A. Schilling. 2010. Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science* **21**(3) 625–642.

Gavetti, G., D. A. Levinthal, W. Ocasio. 2007. Neo-Carnegie: The Carnegie School's past, present, and reconstructing for the future. *Organization Science* **18**(3) 523–536.

Gigerenzer, G., D. G. Goldstein. 1996. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* **103**(4) 650–669.

Grofman, B., G. Owen, eds. 1986. *Information pooling and group decision making*. JAI Press, Greenwich, CT.

Grofman, B., G. Owen, S. L. Feld. 1983. Thirteen theorems in search of the truth. *Theory and Decision* **15**(3) 261–278.

Hambrick, D. C., P. A. Mason. 1984. Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review* **9**(2) 193–206.

Hannan, M. T., J. Freeman. 1977. The population ecology of organizations. *American Journal of Sociology* **82**(5) 929–964.

Hastie, R., T. Kameda. 2005. The robust beauty of majority rules in group decisions. *Psychological Review* **112**(2) 494–508.

Hogarth, R. M. 1978. A note on aggregating opinions. *Organizational Behavior and Human Performance* **21**(1) 40–46.

Hogarth, R. M., N. Karelaia. 2005. Simple models for multiattribute choice with many alternatives: When it does and does not pay to face trade-offs with binary attributes. *Management Science* **51**(12) 1860–1872.

Hogarth, R. M., N. Karelaia. 2007. Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review* **114**(3) 733–758.

Hotelling, H. 1929. Stability in competition. *The Economic Journal* **39**(153) 41–57.

Huber, G. P., R. R. McDaniel. 1986. The decision-making paradigm of organizational design. *Management Science* **32**(5) 572–589.

Janis, I. L. 1972. *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*. Houghton Mifflin, Boston, MA.

Kameda, T., T. Tsukasaki, R. Hastie, N. Berg. 2011. Democracy under uncertainty: The wisdom of crowds and the free-rider problem in group decision making. *Psychological Review* **118**(1) 76–96.

Kanazawa, S. 1998. A brief note on a further refinement of the Condorcet jury theorem for heterogeneous groups. *Mathematical Social Sciences* **35**(1) 69–73.

Knudsen, T., D. A. Levinthal. 2007. Two faces of search: Alternative generation and alternative evaluation. *Organization Science* **18**(1) 39–54.

Kruger, J., D. Dunning. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* **77**(6) 1121–1134.

Laplace, P. S. 1814/1995. *Philosophical Essay on Probabilities*. Springer-Verlag, New York. Translated by A. Dale.

Larrick, R. P., J. B. Soll. 2006. Intuitions about combining opinions: Misappreciation of the averaging principle. *Management Science* **52**(1) 111–127.

Laughlin, P. R. 2011. *Group Problem Solving*. Princeton University Press, Princeton, NJ.

Lawrence, P. R., J. W. Lorsch. 1967. *Organization and Environment: Managing Differentiation and Integration*. Graduate School of Business Administration, Harvard University, Boston, MA.

Liang, D. W., R. Moreland, L. Argote. 1995. Group versus individual training and group-performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin* **21**(4) 384–393.

Mackay, C. 1852. *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds*. Richard Bentley, London.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2** 71–87.

Mezias, J. M., W. H. Starbuck. 2003. Studying the accuracy of managers' perceptions: A research odyssey. *British Journal of Management* **14**(1) 3–17.

Miles, R. E., C. C. Snow. 1978. *Organizational Strategy, Structure, and Process*. McGraw-Hill, New York.

Miller, D., R. Greenwood, R. Prakash. 2009. What happened to organization theory? *Journal of Management Inquiry* **18**(4) 273–279.

Mintzberg, H. 1979. *The Structuring of Organizations*. Prentice-Hall, Englewood Cliffs, NJ.

Mintzberg, H., D. Raisinghani, A. Theoret. 1976. Structure of unstructured decision-processes. *Administrative Science Quarterly* **21**(2) 246–275.

Murray, A. 2010. The end of management. *The Wall Street Journal* Aug 21.

Newell, B. R., D. R. Shanks. 2003. Take the best or look at the rest? factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **29**(1) 53–65.

Nickerson, J. A., T. R. Zenger. 2004. A knowledge-based theory of the firm: The problem-solving perspective. *Organization Science* **15**(6) 617–632.

Payne, J. W., J. R. Bettman, E. J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press, Cambridge, UK.

Radner, R. 1993. The organization of decentralized information processing. *Econometrica* **61**(5) 1109–1146.

Rivkin, J. W., N. Siggelkow. 2003. Balancing search and stability: Interdependencies among elements of organizational design. *Management Science* **49**(3) 290–311.

Sah, R. K., J. E. Stiglitz. 1986. The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review* **76**(4) 716–727.

Schwenk, C. R. 1984. Cognitive simplification processes in strategic decision-making. *Strategic Management Journal* **5**(2) 111–128.

Seshadri, S., Z. Shapira. 2003. The flow of ideas and timing of evaluation as determinants of knowledge creation. *Industrial and Corporate Change* **12**(5) 1099–1124.

Siggelkow, N. 2001. Change in the presence of fit: The rise, the fall, and the renaissance of Liz Claiborne. *Academy of Management Journal* **44**(4) 838–857.

Siggelkow, N., J. W. Rivkin. 2005. Speed and search: Designing organizations for turbulence and complexity. *Organization Science* **16**(2) 101–122.

Simon, H. A. 1947/1997. *Administrative Behavior*. 4th ed. The Free Press, New York.

Simon, H. A. 1956. Rational choice and the structure of the environment. *Psychological Review* **63**(2) 129–138.

Soll, J. B., R. P. Larrick. 2009. Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **35**(3) 780–805.

Sorkin, R. D., C. J. Hays, R. West. 2001. Signal-detection analysis of group decision making. *Psychological Review* **108**(1) 183–203.

Surowiecki, J. 2004. *The Wisdom of Crowds*. Doubleday, New York.

Thompson, J. 1967. *Organizations in Action: Social Science Bases in Administrative Theory*. McGraw-Hill, New York.

Tindale, R. S., T. Kameda, V. B. Hinsz. 2003. Group decision making. M. A. Hogg, J. Cooper, eds., *The SAGE Handbook of Social Psychology*. Sage, London, UK, 381–406.

Tripsas, M., G. Gavetti. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal* **21**(10-11) 1147–1161.

Tushman, M. L., P. Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* **31**(3) 439–465.

Wolfers, J., E. Zitzewitz. 2004. Prediction markets. *Journal of Economic Perspectives* **18**(2) 107–126.

Woodward, J. 1965. *Industrial Organization: Theory and Practice*. Oxford University Press, London, UK.

A Appendix: Screening Function of Each Structure

The respective screening functions for each of the four structures (plus imperfect delegation) are detailed here (these equations are the equivalents of equation (2)). The performance of each structure can be computed by evaluating equation (3) with the corresponding screening function. A detailed log that includes the derivation of the metrics is available from the authors, as is simulation code that verifies these derivations are correct.

A.1 Individual

$S_{\text{ind}}(q, t; e) = 1 - \Phi(0, |t - e|)\big|_{-q}$

A.2 Delegation

$$S_{\text{del}}(q, t; e_L, e_M, e_H) = \begin{cases} S_{\text{ind}}(q, t; e_L) & \text{if } t < \frac{e_M + e_L}{2} \\ S_{\text{ind}}(q, t; e_M) & \text{if } \frac{e_M + e_L}{2} \leq t \leq \frac{e_M + e_H}{2} \\ S_{\text{ind}}(q, t; e_H) & \text{if } t > \frac{e_M + e_H}{2} \end{cases}$$

A.2′ Imperfect Delegation

$$S_{\text{del}'}(q, t, r; e_L, e_M, e_H) = \begin{cases} S_{\text{ind}}(q, t; e_1) & \text{with } \mathbb{P}(1 - \frac{2r}{3}) \\ S_{\text{ind}}(q, t; e_2) & \text{with } \mathbb{P}(\frac{r}{3}) \\ S_{\text{ind}}(q, t; e_3) & \text{with } \mathbb{P}(\frac{r}{3}) \end{cases}$$

Where e_1 is the first-best expert given project t (i.e., the number among e_L, e_M, e_H that is closest to t); e_2 is the second-best expert; and e_3 is the worst expert.

A.3 Voting

$S_{\text{vot}}(q, t; e_L, e_M, e_H) = \mathbb{P}(A, A, R) + \mathbb{P}(A, R, A) + \mathbb{P}(R, A, A) + \mathbb{P}(A, A, A)$

where,

$$\mathbb{P}(v_L, v_M, v_H) = \prod_{i \in \{L,M,H\}} \left[\begin{cases} S_{\text{ind}}(q, t; e_i) & \text{if } v_i = A \text{ (individual } i \text{ accepts)} \\ 1 - S_{\text{ind}}(q, t; e_i) & \text{if } v_i = R \text{ (individual } i \text{ rejects)} \end{cases} \right]$$

A.4 Averaging

$$S_{\text{avg}}(q, t; e_L, e_M, e_H) = 1 - \Phi\left(0, \frac{1}{3}\sqrt{(t - e_L)^2 + (t - e_M)^2 + (t - e_H)^2}\right)\Bigg|_{-q}$$

B Appendix: Robustness Checks

This appendix presents and discusses figures corresponding to the robustness checks mentioned in Section 4.4. The overall conclusion is that the results presented in the body of the paper are qualitatively robust under a broad range of model specifications.

B.1 Randomly Locating the Experts

Figure B.1 plots the performance of structures in which the experts are not located symmetrically (previously, $e_M - e_L = e_H - e_M = \beta$). In this check, the position of each of the three individuals is randomly picked from the range $[e_M - \beta, e_M + \beta]$ (where e_M is no longer the position of the middle expert and instead just the middle of the firm's expertise range). Comparing Figure B.1 to the base case (Figure 2) shows that the ordering of the lines remains unchanged. The performance of the Individual structure is now decreasing in β because its expertise is no longer ideally positioned at e_M.

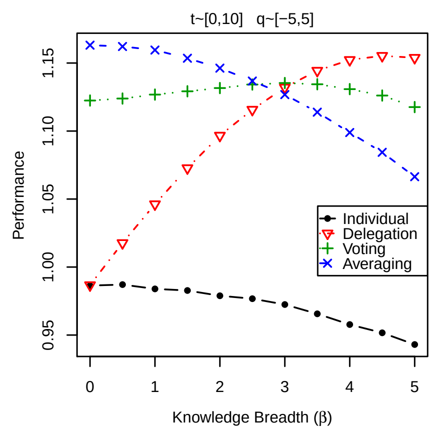


Figure B.1: Randomly locating the decision makers within the range $[e_M - \beta, e_M + \beta]$.

B.2 Adding a Base Error Rate to Individual Perceptions

Equation (1) assumes that the standard deviation of the individuals' perceptions is proportional to $|t - e|$. In this robustness check we explore the case of individuals whose standard deviation does not reach zero even if their expertise perfectly matches the type of the project screened. We therefore add a constant k to the standard deviation of the individuals' perceptions (i.e., the noise in equation (1) becomes $\tilde{n} \sim N(0, |t - e| + k)$). Figure B.2 plots the base case when $k = 0.5$ and $k = 1$ (left and right panels, respectively). The figure shows that, although overall performance decreases (compare left and right panels), the ordering relationship among the different structures' performance remains unchanged.

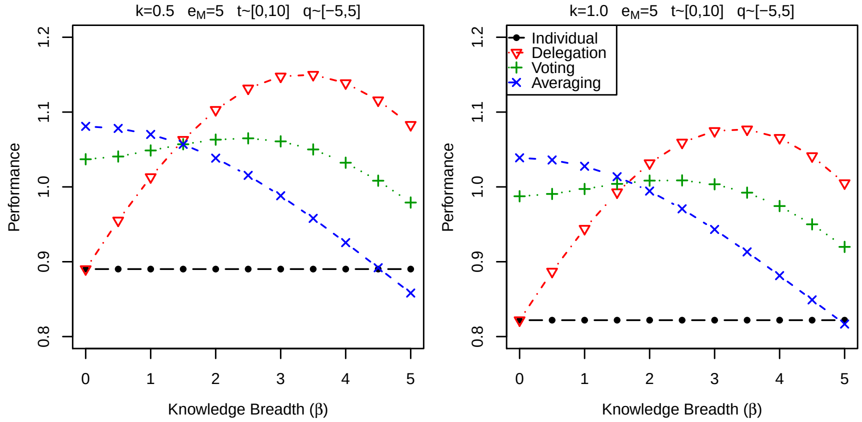


Figure B.2: Adding a base error rate (k) to the individuals' perceptions.

B.3 Scaling the Error in Individual Perceptions

In this robustness check we explore how the results change if the magnitude of the errors is scaled by a factor s (i.e., the noise in equation (1) becomes $\tilde{n} \sim N(0, s|t-e|)$. Figure B.3 shows that the performance of all structures decreases when s increases. The relative ordering of the different structures, however, remains unchanged.

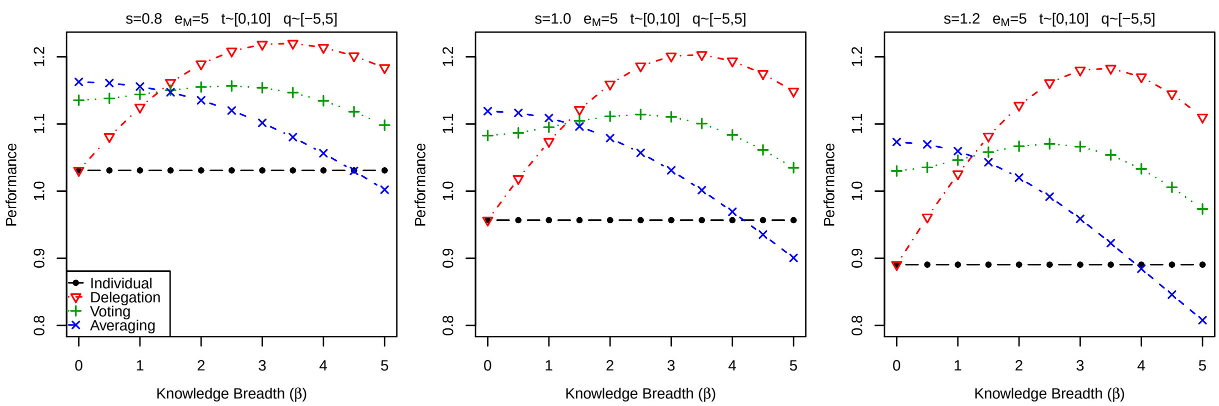


Figure B.3: Scaling the individuals' noise level by factor s.

B.4 Normally Distributed Project Types

In the base case we assumed that the projects screened by the firm are uniformly distributed over $[\underline{t}, \overline{t}]$. In this robustness check we explore a normal distribution of projects. Figure B.4 plots the base case but under $t \sim N(\mu_t, \sigma_t)$ rather than $t \sim U[\underline{t}, \overline{t}]$. Across the three panels, we assume $\mu_t = 5$, and $\sigma_t = 2$, 3, and 4, respectively. Again, the relative ordering of the different structures remains unchanged.

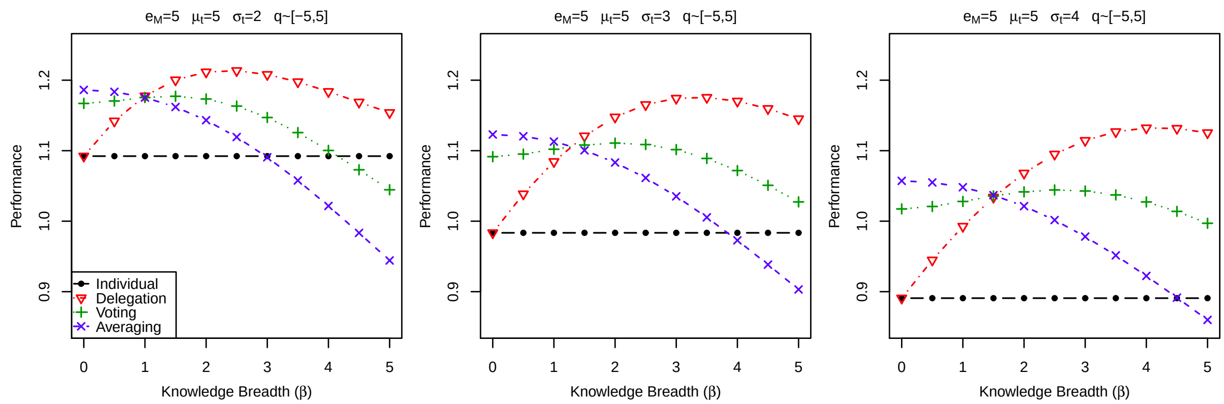


Figure B.4: Performance of the four structures when project type is distributed according to $t \sim N(\mu_t, \sigma_t)$.

C Appendix: Extensions

This appendix provides the mathematical definitions and figures corresponding to the modeling extensions discussed in Section 4.4.

C.1 Correlated Errors

We modeled correlated errors by drawing the individuals' noises from a multivariate normal in which all the cross-correlations are set to ρ. Mathematically, this is implemented by drawing $\tilde{n}_L$, $\tilde{n}_M$, $\tilde{n}_H$ from a multivariate normal with $\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and covariance matrix $\boldsymbol{\Sigma} = \mathbf{D}\boldsymbol{\rho}\mathbf{D}$, where correlation matrix $\boldsymbol{\rho} = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$, and standard deviations matrix $\mathbf{D} = \begin{pmatrix} |t-e_L| & 0 & 0 \\ 0 & |t-e_M| & 0 \\ 0 & 0 & |t-e_H| \end{pmatrix}$.

To analyze the effect of correlated errors, each of the panels in Figure C.1 use the same parameters as in the base case ($e_M = 5$, $q \sim U[-5,5]$, $t \sim U[0,10]$) while varying parameter ρ from 0 (no correlation) to 1 (perfect correlation) in 0.2 increments.

C.2 Optimally Weighted Structures

To create optimally weighted versions of Voting and Averaging in the spirit of Ben-Yashar and Nitzan (1997), we mathematically specified the structures and let a numerical algorithm find the optimal parameters (as a function of β and the environment $[\underline{q}, \overline{q}]$, $[\underline{t}, \overline{t}]$). Note that the closed-form derivations of Ben-Yashar and Nitzan (1997) cannot be directly used in our model, as they only apply to voting and do not take into account project types (i.e., parameters q and t, which are essential to our model). We specified the optimal structures as follows:

- Optimally Weighted Voting only accepts a project if $w_L \mathbb{1}[q'_L > \tau] + w_M \mathbb{1}[q'_M > \tau] + w_H \mathbb{1}[q'_H > \tau]) > \frac{1}{2}$, and
- Optimally Weighted Averaging only accepts a project if $w_L q'_L + w_M q'_M + w_H q'_H > \tau$,

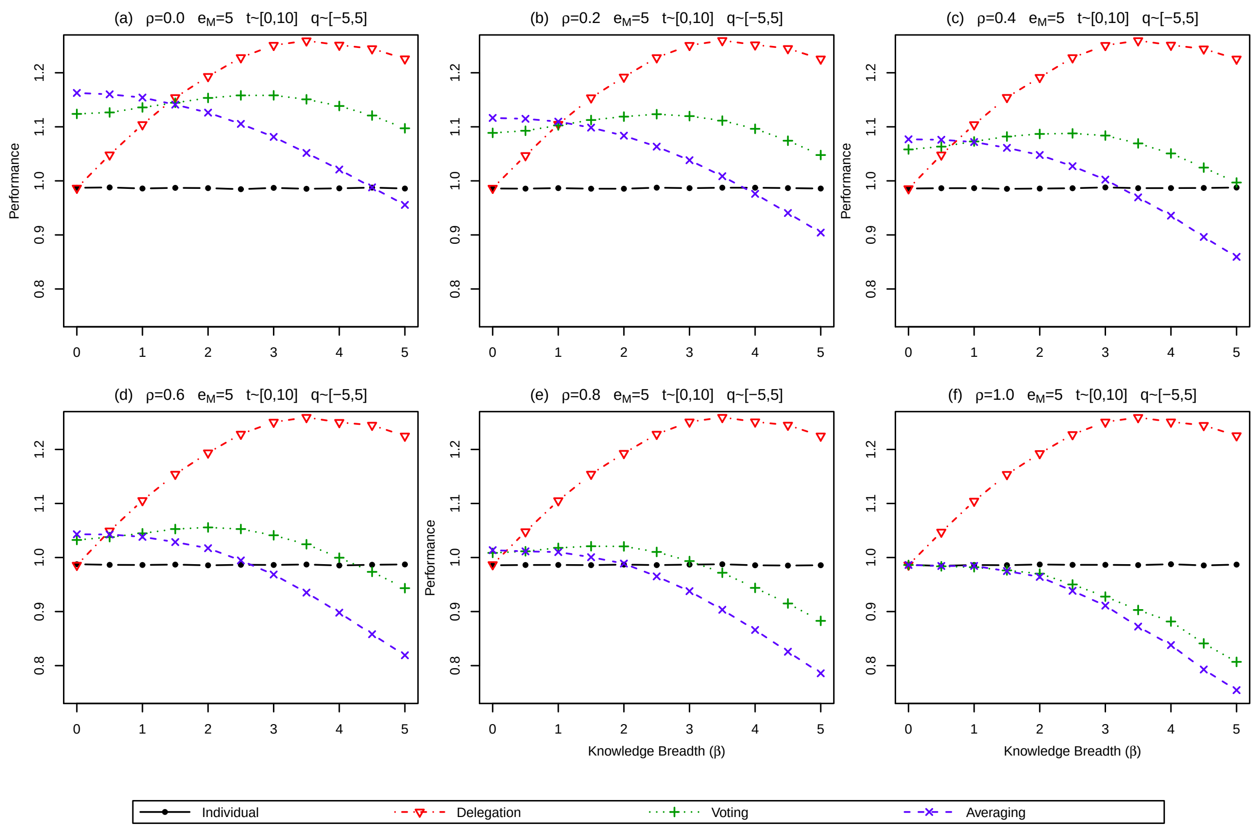


Figure C.1: Performance of the four structures as the correlation (ρ) of the perceptual errors varies from 0 (in the first panel) to 1 (in the last panel).

where w_L, w_M, w_H represent the optimal weights for each individual, and τ is the optimal approval threshold. Optimally Weighted Voting is defined in this fashion in Ben-Yashar and Nitzan (1997:181) and Optimally Weighted Averaging is a natural translation of this idea to the realm of Averaging. Figure C.2 compares Voting and Averaging to their optimal counterparts as well as to Delegation.

C.3 Regression to the Mean of Reported Judgments

In this extension individuals "correct" their judgements before reporting them. The corrected perception is defined as $q'_{\text{corrected}} = f(\cdot)q' + (1 - f(\cdot))\frac{\underline{q}+\overline{q}}{2}$, where $0 \leq f(\cdot) \leq 1$ represents a self-confidence function. We explored two different specifications for $f(\cdot)$:

- Monotonically decreasing. Under this function, if an expert receives a project that perfectly matches her type (i.e., $t = e$) then her self-confidence is 1. As the difference between t and e increases, self-confidence linearly decreases. At the extreme, when the difference between t and e is maximal (i.e., the expert receives the project that it is farther away in the project

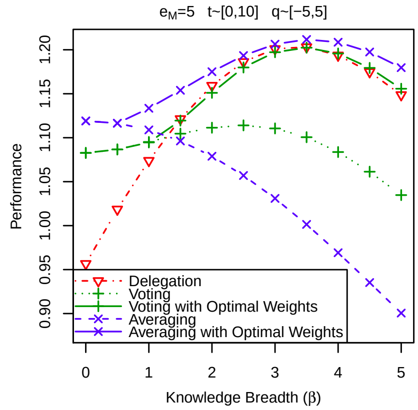
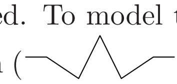


Figure C.2: Performance of Voting and Averaging with and without optimal weights.

range), self-confidence becomes 0. Mathematically, this function is defined as:

$$f_L(t,e) = \min\left(0, 1 - \frac{|t-e|}{\epsilon_{max}}\right), \text{ where } \epsilon_{max} = \max\{|\bar{t}-e|, |\underline{t}-e|\}.$$

- U-shaped. To model the U-shaped self-confidence function we created a "Mexican hat"-like function (), which achieves a maximum when $t = e$, minima at some intermediate points $t = e \pm \lambda(\bar{t} - \underline{t})$, and from there on increases until reaching a plateau level when $t = e \pm (\bar{t} - \underline{t})$. This function is controlled by four parameters: y_1 (level of maximum), y_2 (level of the minima), y_3 (plateau level), and λ (position of the minima as a fraction of the range of possible project types). Mathematically, this function is defined as:

$$f_U(t,e) = \begin{cases} y_3 + \frac{y_2 - y_3}{(e-\lambda(\bar{t}-\underline{t}))-(e-(\bar{t}-\underline{t}))}(t-(e-(\bar{t}-\underline{t}))) & \text{if } e-(\bar{t}-\underline{t}) \leq t < e-\lambda(\bar{t}-\underline{t}) \\ y_2 + \frac{y_1-y_2}{e-(e-\lambda(\bar{t}-\underline{t}))}(t-(e-\lambda(\bar{t}-\underline{t}))) & \text{if } e-\lambda(\bar{t}-\underline{t}) \leq t < e \\ y_1 + \frac{y_2-y_1}{(e+\lambda(\bar{t}-\underline{t}))-e}(t-e) & \text{if } e \leq t < e+\lambda(\bar{t}-\underline{t}) \\ y_2 + \frac{y_3-y_2}{(e+(\bar{t}-\underline{t}))-(e+\lambda(\bar{t}-\underline{t}))}(t-(e+\lambda(\bar{t}-\underline{t}))) & \text{if } e+\lambda(\bar{t}-\underline{t}) \leq t < e+(\bar{t}-\underline{t}) \\ y_3 & \text{otherwise} \end{cases}$$

To study these adjustment processes, the panels in Figure C.3 plot the performance of the four main structures, under f_L and f_U (this last one, assuming $\lambda = 0.25$, $y_1 = 1$, $y_2 = 0$, and $y_3 = 0.8$), using the same assumptions as in the base case ($e_M = 5$, $q \sim U[-5,5]$, $t \sim U[0,10]$).

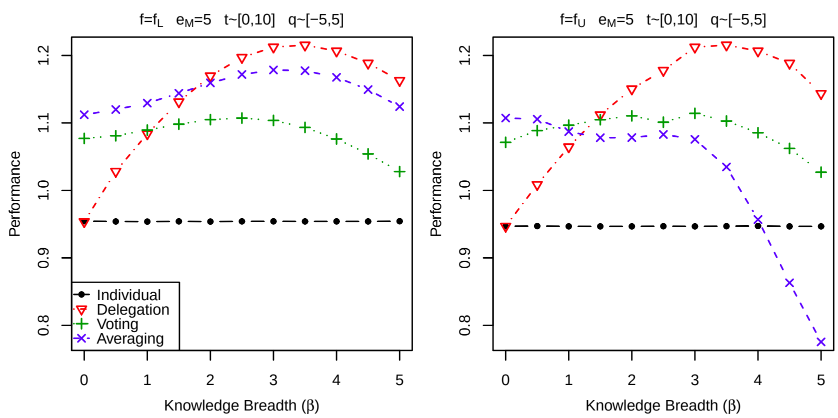


Figure C.3: Performance of the structures under individuals that correct their judgments according to self-confidence. Left: self-confidence is linearly related to accuracy. Right: Self-confidence is U-shaped with respect to accuracy.