Limits to the Wisdom of the Crowd in Idea Selection*

Felipe A. Csaszar
Ross School of Business, U. of Michigan
Ann Arbor, MI 48109
`fcsaszar@umich.edu`

June 25, 2018

Forthcoming in *Advances in Strategic Management*

---

*For helpful comments, the author thanks Allan Afuah and the seminar participants at the University of Michigan and the University of Oxford. The author thanks Jay Zhu and Deepu Nadimpalli for excellent research assistance. The author is grateful for financial support from the 3M Company.

Limits to the Wisdom of the Crowd in Idea Selection

Abstract

An emerging management trend is to use the "wisdom of the crowd" to make decisions traditionally made by the top management alone. Research on this phenomenon has focused mainly on the capacity of crowds to *generate* ideas, but much less is known about a crowd's capacity to *select* ideas. To study crowd-based idea selection in firms, this paper develops a mathematical model of a crowd that makes decisions by majority voting. The model takes into account contingencies that are of particular importance to firms, namely: the size of the population from which the crowd is drawn, the distribution of accuracy among members of the population, and the firm's ability to recruit the population's most accurate individuals. The results show that: (i) under relatively common conditions, increasing the size of the crowd may actually reduce performance; (ii) near-optimal performance can usually be achieved by a much smaller crowd than the one required to achieve optimal performance; (iii) determining the best crowd size depends critically on the firm's ability to recruit "accurate" individuals; and (iv) good performance does not require large crowds unless all population members exhibit low levels of accuracy.

Keywords: information aggregation; organizational structure; organization design; majority voting; crowdsourcing

1 Introduction

An emerging management trend is to use the "wisdom of the crowd" to make decisions that were, traditionally, the sole purview of top management (Surowiecki 2004, Howe 2008). For instance, the T-shirt manufacturer Threadless lets its consumers vote on which products the firm should launch (Lakhani and Kanji 2008), and the mining company Goldcorp published its geological data online and let anyone recommend where to mine next (Tischler 2002). Other firms have tapped "crowds" that are internal to the firm. For instance, HCL Technologies asked its 8,000 employees to review and vote on 300 business plans developed by its managers (Gast and Zanini 2012); also, firms including Microsoft, Google, and HP have used internal prediction markets that allow large numbers of employees to forecast the success of new products (Thompson 2012). Projecting this trend, some commentators predict the "end of management": that firms, in the future, may not need top management teams because firm decisions can be better made by crowds (Murray 2010).

Apart from being an emergent trend, the idea of using crowds to select projects is supported by a classic result in the theory of voting: Condorcet's (1785/1994) jury theorem. This theorem states that if individuals who have a modicum of accuracy—that is, are better than flipping a coin when picking between a good and a bad alternative—make decisions using majority voting, then the accuracy of the group approaches perfection as the number of individuals increases. For example, if each individual had a probability of 70% of picking the right choice, a group of three using majority voting would have a probability of 78%, a group of five would have a probability of 84%, and a group of 29 would have a probability of 99% (how to compute these numbers is discussed later).

But despite its recent popularity and age-old theoretical support, crowds are not commonly used by firms to select projects. Hence, it is natural to wonder whether firms are missing a great opportunity or whether, on the contrary, Condorcet's jury theorem does not apply to real firms.

The goal of this paper is to explore the robustness of the Condorcetian logic when relevant characteristics of real-world firms are taken into account. To do so, this paper draws on and extends two literatures: (a) work in political science on Condorcetian voting and (b) work in the Carnegie tradition on idea selection. The work in political science on Condorcetian voting has studied crowd decision-making but has not paid attention to the specifics of how organizations use crowds to select ideas. Research on Condorcetian voting has studied, for example, the optimality of the weighted majority rule (Nitzan and Paroush 1982, Shapley and Grofman 1984) and the performance of majority voting as the homogeneity assumption is relaxed (Berend and Sapir 2005, Goldstein et al. 2014). In contrast, the research in the Carnegie tradition has studied idea selection in more realistic organizations but has paid little attention to idea selection by large groups. Research here has studied how individual-level biases and incentives affect the quality of the ideas selected (Baumann and Martignoni 2011, Reitzig and Sorenson 2013, Baumann and Stieglitz 2014) and how decision-making structure affects errors of omission and commission in idea selection (Knudsen and Levinthal 2007, Christensen and Knudsen 2010, Csaszar 2012, 2013). The premise of this paper is that it is worthwhile to bridge these two literatures—Condorcetian voting and the Carnegie tradition—to develop a behaviorally plausible model of crowd-based idea selection in firms.

In particular, the model here developed takes into account the population available to the firm (i.e., the size and accuracy distribution of the crowd that the firm can potentially tap into) and the ability of the firm to recruit the population's most accurate individuals. Incorporating the recruiting process into the model is a critical step toward more behavioral realism, as there is much evidence that evaluating expertise and recruiting are noisy processes (see, e.g., Tetlock 2005, Peter and Hull 1969). The model also differs from previous work in that it studies the crowd size that produces satisficing, not optimizing, performance. These modifications lead to results that are novel and meaningful to the literature on organization design.

The main question addressed by this paper is: What is the best size for a crowd that is tasked with making decisions via majority vote? Four novel findings emerge from the analysis. First, there are many cases in which increasing the size of the crowd decreases performance—even if recruiting a larger crowd is costless. Second, near-optimal performance can usually be achieved by a much smaller crowd than the size of one needed to achieve optimal performance. Third, the firm's ability to recruit the most accurate members of the population plays a key moderating role in how the best crowd size is determined. If recruiting ability is high, how many individuals to recruit only depends on the population size and on the ability of the best members of the population; if recruiting ability is low, how many individuals to recruit depends also on the *inability* of the population's least accurate members. Finally, large crowds are needed only when none of the population's members are accurate. These results suggest that there are few situations in which assembling the largest possible crowd is desirable. From a practical standpoint, this paper sheds light on an important organization design question: under what situations and to what extent it is useful to use crowds to select ideas in firms.

2 Theoretical Motivation

Prior work on crowd-based decision making in firms has focused mainly on idea generation—in particular, on how the number and quality of ideas generated by a crowd depend on the characteristics of its individual members (Jeppesen and Lakhani 2010, Bayus 2013), on the incentive scheme adopted (Boudreau et al. 2011, Erat and Krishnan 2012), and on the team process employed (Girotra et al. 2010). Most of this research on idea generation has been done in the context of open innovation and crowdsourcing.[1]

In addition to generating ideas, crowds are also used to select ideas (Bonabeau 2009, Terwiesch and Ulrich 2009, King and Lakhani 2013). However, little is known about idea selection by crowds. Among the studies already cited, only Girotra et al. (2010) have studied

---

[1] Research on open innovation studies how firms can develop new products or services by using ideas that are developed outside the firm boundaries (Chesbrough 2003, King and Lakhani 2013). In turn, research on crowdsourcing focuses on a particular mechanism for achieving open innovation: outsourcing a task to a crowd via an open call such as a competition (Howe 2008, Afuah and Tucci 2012).

crowd-based idea selection in firms. These authors showed by experimental means that a hybrid team process (i.e., one in which individuals first work alone and then together) is better at generating ideas and also at selecting ideas.

Despite the dearth of research, understanding how organizations can use crowds to *select* ideas is important for several reasons. First, without idea selection, even the most comprehensive idea generation effort can prove futile. Second, in many situations the ideas themselves are easy enough to come by ("a dime a dozen," as the adage goes); the difficult aspect is evaluating whether a given idea is a good one. Evaluation is especially problematic for strategic decisions, which are often high-stakes, complex, and nonroutine (Schwenk 1988) as well as hard to demonstrate correct (Laughlin and Ellis 1986). Third, the decreasing costs of information and communication technologies are enabling increasing numbers of firms to use crowds to select ideas; yet, the field of organizations does not offer guidance on when and how to use such crowds.

To fill this gap in the literature, this paper develops a model of crowd-based decision making in firms that draws inspiration and extends two literatures: the Carnegie tradition and Condorcetian voting.

2.1 Relationship to the Carnegie Tradition

An essential observation of the Carnegie tradition is that, even though individuals have limited cognitive capabilities, organizations can overcome these limits by designing appropriate decision-making structures (Simon 1947/1997, pp. 92–93). In the Carnegie tradition, organizations enable humans to achieve goals that no individual alone could achieve. In this sense, the Carnegie tradition is founded on an idea that is not unlike the popular notion that "crowds are wise."

In the quest to understand how organizations can overcome the limits of bounded rationality, the Carnegie tradition has studied such topics as search in complex landscapes, organizational learning, and exploration versus exploitation of knowledge. However, with a few notable exceptions (mentioned below), the effect of different decision-making structures has remained mostly unstudied. In fact, Gavetti et al. (2007) say that this is one of the "forgotten pillars" of the tradition. A possible reason for the lack of such research is that the study of decision-making structure poses serious methodological difficulties: it is hard to make statistically significant claims about the decision-making structure of real-world organizations because few of them are available for observation.

Recent work in the Carnegie tradition has tried to overcome the difficulties of studying decision-making structures by taking three different approaches. These approaches include studying "fruit fly"-type of settings, which involve many small but comparable organizations (Csaszar 2012); developing experiments (Reitzig and Maciejovsky 2015); and analyzing mathematical models of organizations (Knudsen and Levinthal 2007, Christensen and Knudsen 2010, Csaszar 2013, Csaszar and Eggers 2013). Interestingly, these approaches have mostly focused on small organizations (with the exception of Reitzig and Maciejovsky 2015, the previous research has focused on studying groups with six or fewer members).[2] Arguably, the focus on small organizations has been motivated by data availability, analytical tractability, and the fact that many important organizational decisions are made by small groups. An unintended consequence of this focus on small groups is that the Carnegie tradition does not have much to say about how firms can use crowds to make decisions.

The model developed in this paper differs from the previous models in the Carnegie tradition with respect to research question, contingencies, and connections to other literatures. As for the research question, this paper seeks to identify the best crowd size whereas previous papers have sought to identify the effect of other structural elements (such as consensus level or aggregation method used), typically, in the context of smaller organizations. In terms of contingencies, this paper studies effects of the population's size, its accuracy distribution, and the firm's recruiting ability—none of which have been addressed in previous work. Finally, in

---

[2] And even if some of the theoretical work could be extended to make predictions regarding larger organizations (e.g., Knudsen and Levinthal 2007, Csaszar 2013), such predictions would not be satisfactory, as such models assume homogeneous members.

terms of connections to other literatures, this paper draws inspiration from the literature on Condorcetian voting models whereas most previous research has relied on the literatures on committees (Sah and Stiglitz 1988) and on judgment and decision making (Hastie and Kameda 2005).

2.2 Relationship to Condorcetian Voting Models

The current paper builds on the group decision-making literature that descends from Condorcet's (1785/1994) work. In his model of group-based decision making, Condorcet relied on three main assumptions: (i) a group of individuals faces a dichotomous choice between two alternatives, one of which is definitely superior; (ii) individuals are characterized by their *accuracy*—that is, the probability of choosing the superior alternative; and (iii) group "performance" corresponds to the likelihood of the group choosing the superior alternative. Condorcet used these assumptions to model social choice in democracies and to show, in his celebrated jury theorem, that (a) performance increases monotonically with group size and (b) performance tends to unity (i.e., to perfect accuracy) as group size approaches infinity; see Boland (1989) for a modern demonstration. Condorcet used his theorem to argue for the superiority of democracy over other forms of government.

Yet it may well be that Condorcet's assumptions are more applicable to crowd-based decision making in firms than to the social choice settings for which they were originally devised. In problems of social choice, each individual may have different preferences and so it seldom makes sense to speak of the "correct" alternative—and it is this multiplicity of preferences that leads to Arrow's (1951) impossibility theorem for an ideal voting structure. In business settings, however, it is natural to define "correct" as the alternative that maximizes the firm's utility (Marschak and Radner 1972).[3]

---

[3] Additionally, two problems characteristic of social choice settings—herding (Bikhchandani et al. 1992) and insincere voting (Feddersen and Pesendorfer 1997)—are less likely to occur in firm settings. The reasons are that crowd-based decision making in firms is usually simultaneous (which precludes herding) and that Feddersen and Pesendorfer (1997) establish the irrelevance of insincere voting when the population is large and individuals share a utility function.

Condorcet's work was rediscovered by Black (1958), and an important body of work has subsequently been built around it. Some of the main developments in this literature are as follows. Grofman et al. (1983, p. 274) and Ben-Yashar and Nitzan (1997) identify optimal vote weights when accuracy is known; Berend and Paroush (1998) show that the asymptotic prediction of Condorcet's theorem holds in the case of heterogeneous individuals provided the population's average accuracy is better than chance; and Karotkin and Paroush (2003) provide an algorithm to compute the optimal group size while assuming that accuracy is homogeneous but decreases when a new member is added to the group (e.g., because of social loafing).

The model in this paper differs from previous Condorcetian models by simultaneously incorporating three contingencies that are essential to crowd-based decision making in firms: (i) the size of the population from which the crowd is drawn; (ii) the populations' accuracy distribution, and (iii) the firm's ability to recruit the most accurate members of the population. In contrast, most previous research has focused on demonstrating asymptotic results and has ignored the effects of differences in population distribution and recruiting ability. This paper's model differs also in terms of the dependent variable used—here, the group size that produces satisficing, not optimizing, behavior—and the method employed to compute exact results for large, heterogeneous groups.

Overall, the aim of this paper is to explore the phenomenon of crowd-based decision making in firms by combining ideas from the Carnegie tradition and Condorcetian voting models. It is worth noting that Condorcet's assumption—that individuals would try to guess what is best for the firm but, being fallible, would not always get it right—is broadly consistent with the notion of bounded rationality that is central to the Carnegie tradition. The approach used in this paper allows understanding the specific trade-offs that firms face under contingencies that are behaviorally plausible and managerially meaningful. The model also serves as a starting point to explore further questions of crowd-based decision making in firms.

3 Model

The model assumes that the firm faces a dichotomous choice between two alternatives, one of which is better than the other, and that the firm can ask a group to identify the better alternative. The firm then chooses the alternative recommended by the group's majority.[4] The group is recruited from a broader population of individuals who are fallible, so there is only some probability that an individual will choose the correct alternative. This probability, referred to as *individual accuracy*, is not equally distributed in the population. The problem faced by the firm is choosing whom to recruit for the group so as to maximize the odds that the group will choose correctly. The rest of this section describes more formally the elements just outlined.

3.1 The Population

Depending on the firm and the problem asked, the population from which the firm can recruit individuals will vary considerably. For example, the population size available to Threadless when asking about T-shirt designs may be larger than the population size available to HP when asking about new enterprise computer systems. Similarly, the population's accuracy distribution can vary widely with regard to tasks like predicting the sales of a computer system (e.g., if industry veterans are more accurate than the average person); it can also be clumped near 50%—as when the task is predicting tomorrow's direction of the S&P 500 Index. The model accounts for populations of different sizes and accuracy distributions as follows.

The population consists of M individuals. Each individual i in the population $(1 \leq i \leq M)$ is described by her accuracy a_i, which is the probability that she will choose the best alternative.

---

[4] According to Hastie and Kameda (2005) and the references therein, majority voting is the decision rule most frequently used by groups. In light of that prevalence, majority voting is a natural decision rule to study in the context of crowd-based decision making. It can be argued that majority voting is so prevalent because it exhibits several desirable properties. Namely, majority voting is fair (May's 1952 theorem formalizes this property), is fast and frugal (Gigerenzer and Goldstein 1996), is robust under different environments and decision makers (Hastie and Kameda 2005, Csaszar and Eggers 2013), and can jointly minimize the chances that errors of omission or of commission will occur (Sah and Stiglitz 1988, Csaszar 2013).

For instance, if $a_i = 0.9$ then individual i will make the correct choice 90% of the time; if $a_i = 0.5$, then she will be no better than chance at predicting the right alternative. Accuracy can range between 1 (always making a perfect choice) and 0.5 (randomly picking right or wrong).[5]

The populations' accuracy is evenly distributed from $a_{\max}$ (the most accurate individual) to $a_{\min}$ (the least accurate individual). For instance, if $a_{\max} = 0.9$, $a_{\min} = 0.5$, and $M = 5$, then the population would consist of individuals $a_1 = 0.9$, $a_2 = 0.8$, $a_3 = 0.7$, $a_4 = 0.6$, and $a_5 = 0.5$.[6] The vote of individual i is denoted v_i, which takes the value of 1 if she made the right choice (or 0 otherwise). Thus, each individual vote is distributed according to $v_i \sim \text{Bernoulli}(a_i)$.

3.2 The Group

The set of N individuals whose opinion is consulted is called the group or crowd.[7] Because the group is recruited from the population, group size (N) can take values between 1 and M (inclusive).

Which N individuals of the population are recruited depends on the firm's recruiting ability. The base model (analyzed in the first two subsections of the Results section) assumes perfect recruiting ability: that the N recruited individuals are the N most accurate members of the population. So in the previous $M = 5$ example, if the organization designer sets $N = 3$ then the individuals with accuracies 0.9, 0.8, and 0.7 would be recruited. The third subsection of the Results relaxes the perfect recruiting assumption by introducing recruiting errors as a contingency.

---

[5] As is common in the Condorcetian voting literature, the model assumes a dichotomous choice. The lower bound of accuracy is set to 0.5, as this is the performance of an uninformed individual selecting between two alternatives.

[6] Further work could explore other distributions of accuracy, such as a bimodal distribution, in which there are a few high-accuracy experts and plenty of low-accuracy individuals. Although the current paper does not study such distributions explicitly, comparing across scenarios can shed light on the performance of different subgroups within a larger population (e.g., comparing using a few high-accuracy individuals versus many low-accuracy ones).

[7] The word "crowd" connotes a large number of individuals. Because the number of consulted individuals can be either large or small, in what follows the more neutral word "group" is favored.

The group's opinion is determined by majority voting. Thus, the group chooses the best alternative when the majority of the group votes for it—that is, if $\sum_{i=1}^{N} v_i > (N-1)/2$. As is customary in the voting literature (see, e.g., Grofman et al. 1983), the only groups analyzed are those with an odd number of members.[8]

3.3 Choosing the Best Group Size

The main question addressed in this paper is: How large should the recruited group be? More specifically, this paper seeks to identify the best value of N as a function of the population size (M), the population's accuracy distribution ($a_{\max}$, $a_{\min}$), and recruiting errors (introduced in the third subsection of the Results).

This paper defines "best" group size not in terms of the *optimal* N (i.e., the N^* that maximizes performance) but rather in terms of a *satisficing* N (i.e., an N^{**} yielding performance that is good enough for most practical purposes; N^{**} is precisely defined in the second subsection of the Results).

Establishing the best N as a function of the model's contingencies amounts to determining which is better: the wisdom of the crowd ($N = M$), the wisdom of one individual ($N = 1$), or some intermediate solution. As shown in the Results section, the answer to this question depends in complex ways on the model's contingencies.

In order to identify the best N, one must compute the probability that a group will make the right decision. Given that each vote follows a Bernoulli distribution, the group vote will follow a Poisson-binomial distribution with parameters $a_1, \ldots, a_N$. Hence *group performance* is the probability that the group will choose the correct alternative and can be written as

$$\text{Performance} = 1 - F(a_1, \ldots, a_N)\big|_{(N-1)/2}, \tag{1}$$

---

[8] Majority voting in groups with an even number of members is seldom analyzed because in these groups ties must be decided in an arbitrary, non-neutral way (e.g., if in a group of eight, four reject a proposal, the proposal will be rejected; yet, if the same four accept the proposal, the proposal will not be accepted). This arbitrariness makes even-groups less likely to make correct decisions than the adjacent (odd) groups (Dougherty and Edward 2009).

where $F(\cdot)$ is the cumulative distribution of the Poisson-binomial.[9]

One difficulty in evaluating Equation (1) is that the computation of $F(\cdot)$ usually involves enumerating a vast number of cases, which is impractical even for small values of N. For instance, $N = 41$ (which is hardly representative of a large crowd) generates over a trillion $\left(= \sum_{i=21}^{41} \binom{41}{i}\right)$ cases. An alternative way to compute $F(\cdot)$ might be via an approximation method. Regrettably, however, the existing approximation methods are not applicable here. Le Cam's approximation would work only if the a_i's tended to 0, and approximations based on the normal distribution do not provide any guarantees on the accuracy of the main variable of interest here: the best value of N.[10]

It is fortunate that efficient methods for the exact evaluation of the Poisson-binomial have been developed in recent years (Fernández and Williams 2010, Hong 2013a).[11] The ensuing results are generated by evaluating Equation (1) under different model assumptions. Numerical analysis is used because the mathematical expressions involved are not analytically tractable.

4 Results

To understand the mechanisms driving these results, the first two subsections explore the model under the assumption of perfect recruiting (i.e., that the most accurate individuals are recruited first). That assumption is then relaxed in the third subsection.

---

[9] For the sake of parsimony, this model ignores the cost of hiring the group members. It would be straightforward to incorporate hiring costs—for example, performance could be redefined as the probability of making the right decision *minus* a cost that is proportional to N—but doing so would add little insight to the results. Moreover, there are many settings in which the group's members work for free.

[10] Numerical experiments show that the best value of N predicted by the normal approximation can differ by more than 100% from the real best value of N. For references to the cited approximations methods, see Hong (2013a).

[11] The results reported in the Results section are computed using algorithm DFT-CF discussed in Hong (2013a) and implemented in Hong (2013b).

4.1 The Determinants of Performance

This section uses Figure 1 as an aid to explicating the effects of group size (N), population size (M), and population accuracy ($a_{\max}$, $a_{\min}$) on performance. Each panel in this figure shows how group performance (on the y-axes) varies as a function of group size (on the x-axes). From left to right, the columns of panels represent small, medium, and large populations ($M = 21$, 101, and 1,001, respectively). From top to bottom, the rows of panels represent populations with low accuracy ($a_{\max} = 0.6$ and $a_{\min} = 0.5$), broad accuracy ($a_{\max} = 0.9$ and $a_{\min} = 0.5$), and high accuracy ($a_{\max} = 0.9$ and $a_{\min} = 0.8$). In each panel, a solid circle (●) marks where maximum performance is achieved (the open circles are explained later). The group size that achieves the maximum performance is called the *optimal group size* and is denoted N^*.

Because the analyses in this section assume that the firm can recruit the most accurate individuals, every curve in Figure 1 starts at $a_{\max}$. The reason is that, when $N = 1$, the group consists of the most accurate individual in the population.

A first observation from Figure 1 is that the effect of population size is most relevant when the population is characterized by low levels of accuracy. Thus, in the first row of panels, maximum performance takes values of roughly 0.7, 0.85, and 1 in panels (a), (b), and (c), respectively; in contrast, maximum performance in the remaining rows of panels is always close to 1. That is to say, using a large group is especially beneficial when the group is ignorant.

In contrast, if the group includes some highly accurate individuals (as occurs in the second and third rows of panels) then—after a few of these individuals have been recruited—the group achieves near-perfect performance. This happens because, once a group includes a core of highly accurate members (whose votes are likely to support the correct alternative), the less accurate individuals (whose votes will be more random) are unlikely to overturn the votes of that accurate core.

A second observation from Figure 1 is that recruiting the entire population is not always

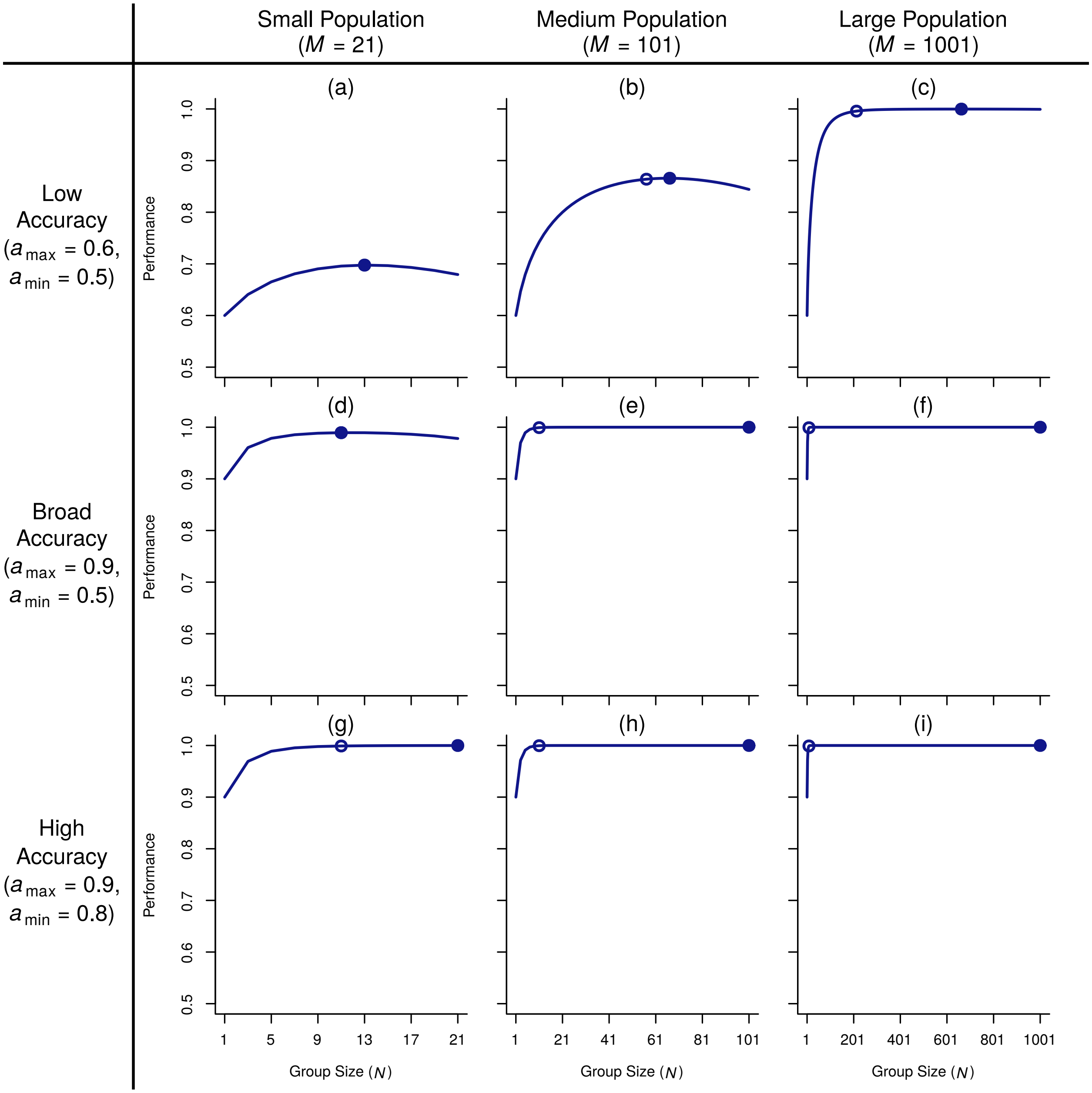

Figure 1: Group performance as a function of group size N (x-axes), population size M (columns), and the accuracy distribution ($a_{\max}$, $a_{\min}$) of the population (rows). In each panel, maximum performance is marked by a solid circle (●) and near-optimal performance (defined in the second subsection of the Results) is marked by an open circle (○).

a good approach. The population is $M = 21$ in both panels (a) and (g); however, only in panel (g) it is optimal to hire everyone (i.e., the $\bullet$ is at the end of the curve) whereas in panel (a) the optimal number to hire is just $N = 13$ individuals. So even though the relationship between group size and performance is monotonically increasing in some cases, in other cases that relation has an inverted U-shape. In Figure 1, panels (a)–(d) exhibit inverted U-shapes while panels (e)–(i) exhibit monotonically increasing performance. This observation qualifies the observation from previous research that smaller crowds can sometimes be better than larger crowds (Goldstein et al. 2014). The necessary condition for smaller crowds outperforming larger ones is not simply that "smarter crowds can be identified in advance" Goldstein et al. (2014, p. 471), as in all panels in Figure 1 the organization designer can perfectly identify who is "smarter" (i.e., the most accurate individuals are recruited first), but not all panels exhibit an inverted U-shape.

This difference in the shape of the performance curves plotted in Figure 1 is explained by a quantity–quality trade-off. On the one hand, increasing the group size can be beneficial: as group size increases, vote counts become less random and more representative of the underlying probabilities (which favor the correct decision, since $a_i > 0.5$). On the other hand, increasing group size could be detrimental: given that recruiting proceeds in decreasing order of accuracy, each additional member is less accurate than those already in the group. In other words, group performance increases with group size only if the marginal benefit from increasing group size is greater than the marginal cost of decreasing group accuracy; and vice versa for the cases where group performance decreases with group size. Numerical methods can be used to determine which of these two marginal effects dominates in any given case (as explained in the Model section, the Poisson-binomial distribution does not yield analytically tractable expressions).

The observation that performance can decline when a larger fraction of the population is recruited casts doubt on the extent to which the "wisdom of the crowd" doctrine (Surowiecki 2004) is valid. Faced with environments like the ones in panels (a)–(d), organizations would be ill advised to hire the largest possible crowd.

This observation also serves as a reminder that Condorcet's (1785/1994) jury theorem—which commonly serves as a theoretical foundation for advocating the wisdom of the crowd—depends on premises that are not generally valid. Condorcet's theorem states that, if individuals are homogeneous ($a_1 = a_2 = \cdots = a_N = a$) and have a modicum of accuracy ($a > 0.5$), then: (i) group performance increases with N (for N odd) and (ii) group performance tends to 1 as group size goes to infinity. Yet if Condorcet's jury theorem applied to the model presented here, then the inverted U-shapes plotted in panels (a)–(d) of Figure 1 would be impossible. Of course, Condorcet's theorem does not apply to the current model because the recruited individuals are not homogeneous and the group size cannot grow indefinitely; instead, this model presumes decreasing accuracy among subsequently recruited individuals as well as a bounded group size. Because the assumptions in the current model are plausible characteristics of real-world organizations and because these assumptions lead to results that are fundamentally different from the Condorcet's jury theorem predictions, it follows that the theorem should be used with caution when the goal is to derive real-world implications.

A final observation from Figure 1 is that near-optimal performance can almost always be achieved by using a much smaller group than what is required to achieve optimal performance. In panel (c), for instance, optimal performance occurs at about $N = 700$ but the performance for $N = 200$ is nearly indistinguishable from optimal. In other words, the firm could forgo recruiting 500 more members and still attain essentially the same performance. Panels (f) and (i) show even starker differences between the group sizes required to achieve optimal versus near-optimal performance.

4.2 The Determinants of Best Group Size

In practice, it is not economical to hire a much larger group to achieve a small improvement in performance. Therefore, the analysis to follow will focus on the group size required to achieve *near*-optimal performance, which is defined as being in the top $X\%$ of the performance range.

For the sake of simplicity, this paper assumes $X = 1$ (i.e., if maximum performance is 100%, then a performance of 99% is near-optimal). As it will be shown later, using 1% as the threshold to define near optimal performance, is enough for the group sizes necessary to attain optimal and near-optimal performance to be vastly different. Although not the focus of this paper, the effect of varying X can be visually appreciated from the figures that display performance as a function of group size (i.e., figures 1 and 5). In practice, X depends on specifics of the situation that determine the cost of increasing the crowd (e.g., wages) vis-à-vis possible gains from increased accuracy (e.g., benefits from selecting good projects).

The smallest group size that achieves this near-optimal performance is called the *best group size* and is denoted N^{**}.[12] The open circles (○) in Figure 1 mark the position of N^{**}. By definition, the best group size cannot be larger than the optimal group size (i.e., $N^{**} \leq N^{*}$).

To explain how the best group size N^{**} depends on the model's parameters, the analysis now focuses on Figure 2. Each panel in this figure shows how N^{**} (on the vertical axes) depends on $a_{\max}$ and $a_{\min}$ (on the horizontal axes). From left to right, the panels represent small, medium, and large populations ($M = 21$, 101, and 1,001, respectively). Because the population's accuracy is distributed from $a_{\max}$ to $a_{\min}$, the response surfaces exist only in those areas where $a_{\max} \geq a_{\min}$.

A first observation from Figure 2 is that, in each panel, the best group size N^{**} increases as $a_{\max}$ approaches 0.5. In other words, the fewer high-accuracy members in the population, the larger the group size required to achieve good performance. This result makes intuitive sense: as the population becomes less accurate, the group can compensate for its less accurate members by recruiting more members.

A second observation from comparing across panels in Figure 2 concerns the fraction of the population that must be recruited in order to achieve near-optimal performance (i.e., N^{**} as

---

[12] Formally, $N^{**} = \min\{N \mid \pi(N) \geq 0.99(\pi_{\max} - \pi_{\min}) + \pi_{\min}\}$, where $\pi(N)$ is performance (defined in Equation 1), $\pi_{\max}$ is the optimal group performance (which occurs at $N = N^{*}$), and $\pi_{\min}$ is the lowest group performance (which occurs at $N = 1$).

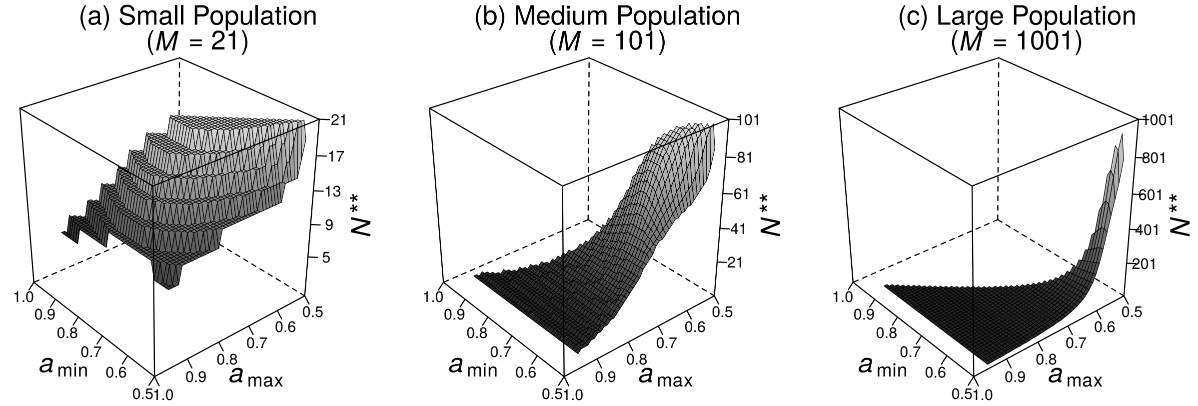

Figure 2: The best group size (N^{**}) as a function of population size (M) and of the population's accuracy distribution ($a_{\max}$, $a_{\min}$).

a fraction of M). For instance, N^{**}/M never falls below 33% (= 7/21) in panel (a) whereas, in panel (c), N^{**}/M is almost always well below that mark. This result can be explained as follows. When there are high-accuracy members in the population (i.e., with $a_{\max}$ close to 1), a handful of individuals can achieve near-optimal performance and they will constitute only a small fraction of a large population (i.e., in panels (b) and (c) but not in panel (a)). But if everyone in the population has low accuracy (i.e., $a_{\max}$ close to 0.5), then each new recruit has only a small effect on performance and so a large fraction of the population is necessary to achieve near-optimal performance—irrespective of the total population size.

To gain a better understanding of how the model parameters affect what fraction of the population to recruit, it is useful to inspect a formula that predicts N^{**}/M as a function of those parameters. Because this is not a closed-form model, one must fit a formula from the predictions of the model. One such formula is:[13]

$$\text{Predicted}\log(N^{**}/M) = 9.00 - 8.11a_{\max} - 0.14a_{\min} - 0.89\log(M). \tag{2}$$

This formula was selected because it provides a good fit ($R^2 = 0.97$) and is parsimonious. The quality of the fit is visible in Figure 3, which plots observed versus predicted values.

---

[13] This formula was estimated using ordinary least-squares regression on a data set that contained the value of N^{**} for all 11,250 combinations of M in $\{3, 7, 11, \ldots, 999\}$ and for $a_{\min}$, $a_{\max}$ in $\{0.50, 0.55, 0.60, \ldots, 0.95\}$ subject to $a_{\min} < a_{\max}$.

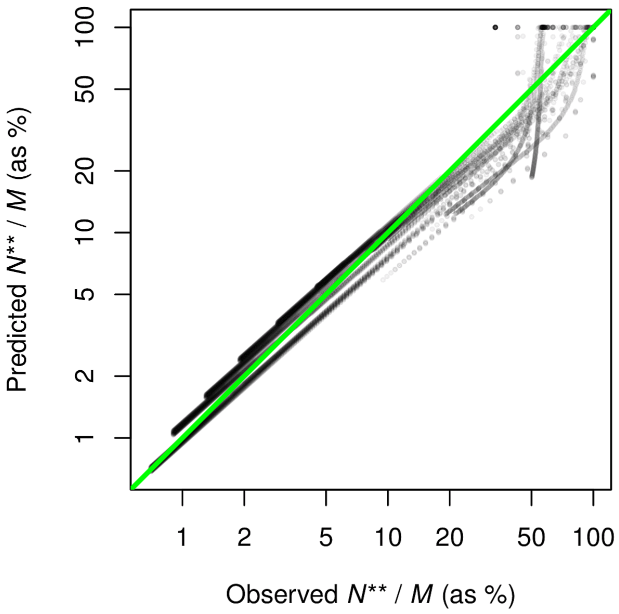

Figure 3: Fit of the formula used to predict what fraction of the population must be recruited in order to achieve near-optimal performance (N^{**}/M). The diagonal line (predicted = observed) represents the perfect fit.

Equation (2) indicates that the fraction of the population that should be recruited in order to achieve near-optimal performance decreases as $a_{\max}$, $a_{\min}$, or M increases. Equivalently, it shows that one should recruit a larger fraction of the population as the population's accuracy (or size) decreases.

An interesting observation from Equation (2) is that the upper end of the population's accuracy distribution matters more than the lower end—the magnitude of the coefficient for $a_{\max}$ is more than 50 times larger than that for $a_{\min}$. This is because the model analyzed so far assumes perfect recruiting, which in turn implies that the most accurate individuals are recruited first; hence the firm rarely needs to scrape the "bottom of the barrel." As a result, $a_{\max}$ is more representative of the typical group member than is $a_{\min}$. In the next section, the assumption of perfect recruiting is relaxed.

4.3 The Effect of Imperfect Recruiting

So far the model has assumed that the firm recruits the N most accurate members of the population. However, such a scenario rarely happens. In the real world, multiple reasons may hamper the firm's ability to recruit in order of accuracy. For instance, individual accuracy levels are seldom perfectly observable; this problem is more acute when the decision concerns a novel problem (i.e., one for which there can be no track record of individuals' accuracy).[14] Perhaps some of the best experts are busy, which forces the firm to recruit instead lower-accuracy individuals. Finally, there may be reasons beyond accuracy for including some individuals in the group. For example, some board members may not be the most accurate decision makers yet nonetheless could enhance group outcomes through their valuable skills or connections.

To account for such cases, the model is extended to allow for a "tunable" degree of recruiting errors. Such errors are controlled by parameter r ($0 \leq r \leq 1$): if $r = 0$, then there are no recruiting errors and the firm recruits (as in the base model) the top N individuals of the population; if $r = 1$, then recruiting is completely random and the firm recruits N random individuals from the population. As r increases from 0 to 1, recruiting becomes increasingly random. Figure 4 illustrates the effect of recruiting errors r on the order in which the population is recruited. The effect of r is implemented via simulation as follows: (i) the population (of size M and of accuracy ranging from $a_{\max}$ to $a_{\min}$) is listed in order of descending accuracy; (ii) a fraction r of randomly chosen individuals from the list is selected; (iii) the selected individuals are randomly shuffled among themselves; and (iv) the firm recruits individuals in positions $1, \ldots, N$ of the resulting list.

Figure 5 is similar to Figure 1, but each panel has multiple curves to account for different degrees of recruiting errors (r). Each curve represents the average performance across 1,000 simulations for a given level of recruiting errors, where r takes the values 0, 0.25, 0.5, 0.75,

---

[14] The difficulty of observing individual accuracy also limits the ability of firms to use decision-making structures that rely on precise information about individual accuracy (such as optimally weighted majority voting; Ben-Yashar and Nitzan 1997).

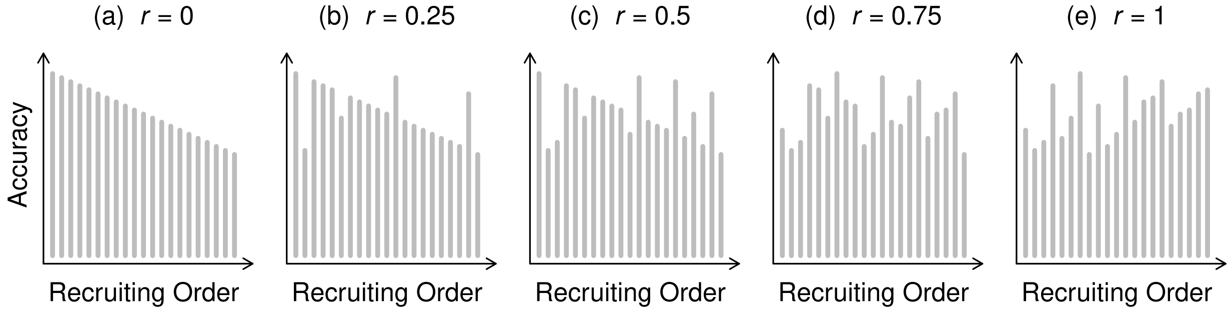

Figure 4: Illustration of the effect of recruiting errors (r) on the order in which the firm recruits individuals.

and 1. Since $r = 0$ yields perfect recruiting, the $r = 0$ curves in Figure 5 are identical to the curves in Figure 1.

In every panel in Figure 5, the performance curves start separate when $N = 1$ and converge to a unique performance level when $N = M$. The starting point of the curves corresponds to the expected accuracy of the first recruited individual. For example, if recruiting is absolutely imperfect ($r = 1$) then all population members are equally likely to be recruited; hence performance starts at the average between $a_{\max}$ and $a_{\min}$ (e.g., in panel (d) the $r = 1$ curve starts at $0.7 = (0.9 + 0.5)/2$). The curves converge at the rightmost end of each panel because recruitment in these cases is not affected by r (since everyone is recruited).

A general observation from Figure 5 is that the effect of recruiting errors r on performance is negative. However, this negative effect varies from considerable (as in panels (a)–(e)) to almost imperceptible (as in panels (f)–(i)).

The effect of imperfect recruiting on performance is most detrimental when the population includes low-accuracy members (i.e., if $a_{\min}$ is low, as in the first and second rows of panels in Figure 5). In these cases, a recruiting error can be costly because it may lead to recruiting an extremely inaccurate group member. In contrast, when the population consists almost solely of high-accuracy individuals (as in the third row of panels in Figure 5), recruiting errors are innocuous: the worst possible error is hiring someone who is still very accurate.

The effect of recruiting errors can sometimes be offset by hiring more individuals. In panel (c), for instance, all the curves achieve maximal performance but there are vast

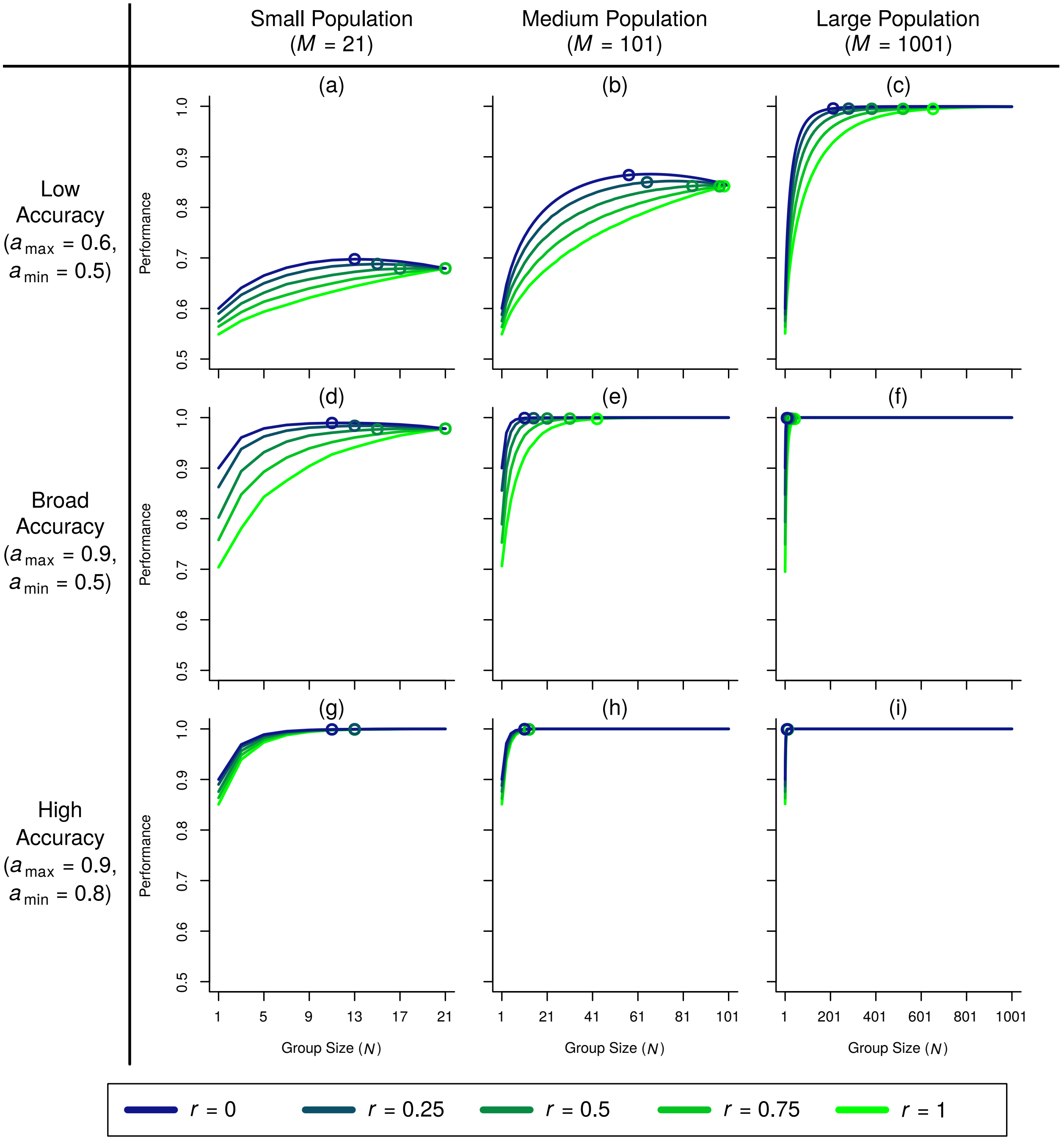

Figure 5: Group performance as a function of recruiting errors r (colored curves), group size N (x-axes), population size M (columns), and the accuracy distribution ($a_{\max}$, $a_{\min}$) of the population (rows). In each panel, near-optimal performance is marked by an open circle (○).

differences in the group size needed to achieve that performance. If recruiting is perfect ($r = 0$) then a group of about 200 achieves near-optimal performance; in contrast, if recruiting is absolutely imperfect ($r = 1$) then a group of about 650 is needed to achieve the same performance level. In cases like this, the gist of Condorcet's theorem applies: if individuals who are any more accurate than a coin flip are recruited in large enough numbers, then together they can yield performance that is almost perfect.[15]

However, there are other cases where larger groups cannot compensate for recruiting errors. In panel (b), the maximum performance achieved by the $r = 0$ curve (at around $N = 60$) is not matched by any other curve in that panel. The cases where larger groups cannot compensate for recruiting errors are those in which performance under perfect recruiting follows an inverted U-shape (the cause of this shape was discussed in the first subsection of the Results). Under these circumstances, recruiting ability (i.e., achieving a low r) can provide a competitive advantage that cannot be matched by recruiting a larger group.

To understand how recruiting errors r affect the fraction N^{**}/M of the population that should be recruited to achieve near-optimal performance, it is helpful to look at a fitted formula comparable to that presented for the case of perfect recruiting (Equation 2), but under absolutely imperfect recruiting ($r = 1$). The fitted formula in this case is:[16]

$$\text{Predicted} \log(N^{**}/M) = 9.50 - 5.68 a_{\max} - 3.84 a_{\min} - 0.80 \log(M). \tag{3}$$

Like Equation (2), this formula yields a good fit to the data ($R^2 = 0.96$).

The main difference between this formula and Equation (2) is that now the coefficient for $a_{\min}$ has a much larger magnitude. Recall that the coefficient for $a_{\min}$ was close to 0 in Equation (2). Now, both ends of the population—and not just $a_{\max}$—matter for determining what fraction of the population should be recruited. In other words, the firm's ability to

---

[15] This observation is consistent with extensions of Condorcet's jury theorem to heterogeneous groups (Berend and Sapir 2005).

[16] This formula was estimated on a data set identical to the one described in note 13; the only difference here is that the dependent variable results from averaging 1,000 simulations under the condition $r = 1$.

recruit the best members of the population plays a key moderating role in how the best group size is determined. If recruiting ability is high ($r = 0$), then the ideal number of recruits depends on the population's size and on the abilities of its most accurate members. If recruiting ability is low ($r = 1$), then the number of individuals to recruit depends also on the level of inaccuracy exhibited by the population's "worst" members.

5 Discussion

This study introduces a mathematical model that predicts how well a group performs at an idea selection task. The model extends previous Condorcetian voting models by including parameters that are critical in crowd-based decision making by firms (namely, the size of the population from which the crowd is drawn, the population's accuracy distribution, and the firm's ability to recruit the population's most accurate individuals) and by analyzing performance metrics that are meaningful for firms (i.e., the group size needed to achieve near-optimal performance). The results demonstrate that: (i) there are many cases in which increasing the crowd size decreases performance; (ii) near-optimal performance can be achieved by a group that is much smaller than the size of one needed to achieve optimal performance; (iii) determining the best crowd size depends critically on the firm's ability to recruit accurate individuals; and (iv) large crowds are needed only when everyone in the population is of low accuracy.

5.1 Organization Design Implications

As the cost of information and communication technologies continues to decrease, it will become easier to use crowds to make decisions. Given this availability, it is important that organization designers know when and how to use crowds to make decisions.

To this respect, a first distinction organization designers should keep in mind is the difference between using crowds to generate and to select ideas. Using crowds to generate ideas can be understood in terms of distant search and the access to novel and diverse ideas it provides (Afuah and Tucci 2012); in contrast, using crowds to select ideas can be understood in terms of recruiting voters with different accuracies (as elaborated in the current paper). Distant search and voting are different processes that call for different considerations; hence, organization designers should treat both stages differently.

In practical terms, this paper suggests that using very large crowds to select ideas will be rarely necessary (only when everybody in the population has very low accuracy). In contrast, many times much smaller groups will do as well or better than larger crowds. For instance, smaller groups (or even just one individual) are better when the firm has access to superior experts (e.g., as shown in Figure 1(b), a crowd of one hundred individuals whose accuracy is in the 0.5–0.6 range would not outperform one expert of accuracy 0.9). Smaller crowds are also better than larger ones when recruiting errors are low (i.e., low r curves in Figure 5). And, many times, the performance improvement of increasing crowd size is almost negligible (e.g., in most of Figure 1 near-optimal performance is achieved with a fraction of the members required for achieving optimal performance). In light of these findings, Murray's (2010) prognostication about the "end of management" seems unlikely. Another implication for organization design is that because the optimal crowd size decreases with recruiting ability, it may be worthwhile for firms to invest in improving this ability.

Apart from the considerations developed in this article, there are additional reasons to be conservative about the extent to which crowds will be used to replace traditional management. For instance, crowds may not be useful if the cost of leaking information is high (e.g., a firm may want to keep a strategy secret for as long as possible). Also, using crowds may not be cost-effective if motivating the crowd to participate is expensive either in terms of payments necessary to recruit the crowd or in terms of time (e.g., hours that internal crowd members will spend deliberating among choices). Finally, crowds may be impractical if communicating enough details to the crowd is cumbersome (e.g., explaining enough about a new product, so that the crowd members can emit informed judgments).

Taking all these considerations into account reduces the set of cases where crowds can be useful. It is in this reduced set, where organization designers should consider carefully whether or not to use crowds to select ideas. For instance, if no one is very accurate at making a certain type of choice and if using a crowd does not risk leaking information and is not prohibitively expensive or slow, then using a crowd could make sense.

5.2 Limitations and Further Work

As all research, this study has limitations, which could be addressed by future empirical and theoretical work. In terms of empirical work, a first extension is to test the relationships hypothesized in this paper and study the extent to which other process characteristics (such opportunities for communication and deliberation) affect the group's performance in selecting ideas. Another way to extend the current work is to examine what is the accuracy distribution from which firms hire and how it depends on the problem at hand. Most likely different problems exhibit vastly different distributions (e.g., in some situations, experts may be much more accurate than others, producing skewed or bimodal distributions). A benefit of knowing the accuracy distribution available to a given firm is enabling a more precise estimation of the best crowd size. Another avenue for empirical research is studying how crowd-based decision making affects firm performance in ways other than influencing the choice of the best alternative. For instance, using crowds may increase employee buy-in and improve brand recognition; however, it might also compromise the firm's intellectual property. Further empirical work is needed to understand the significance of such effects vis-à-vis the ones studied in this paper.

In terms of theoretical work, further research could compare the performance of voting to that of other idea selection mechanisms that can be used by crowds, such as averaging predictions and prediction markets. Further work could also study selection among more than two choices as well as selection of choices that are not independent (e.g., picking product attributes that interact among themselves, such as price, quality, and location).

Another promising area of research is to understand how incentive misalignments affect crowd decisions. The current model assumes that all individuals are "well-intended" and thereby make predictions trying to choose what is best for the firm. But such degree of incentive alignment may not be realistic. Hence, another avenue for increased realism is exploring how incentives affect crowd design. A possible application is devising weighting schemes to reduce biases introduced by incentive misalignments.

5.3 Conclusion

New communication technologies have made possible—for the first time in the history of the firm—to open the decision making of firms on a large scale. Today it is much more difficult to determine how large a decision-making group should be than it was a few years back; this is because most firms can now assemble groups of essentially any size, from one individual to the whole population. A popular answer to the group size question is one that relies on the "wisdom of the crowd." Arguably, this answer is popular because it evokes the legitimacy of democracy and of the mathematical arguments supporting it (e.g., Condorcet 1785/1994, Galton 1907). Even so, most firms do *not* use large crowds to make decisions. Hence a practical puzzle is whether firms should or should not embrace the idea of using crowds to make decisions.

The purpose of this research has been to study in a rigorous manner the question of whether firm decisions should be made by small or large groups while taking into account aspects of the problem that are relevant to firms, such as the population of from which the crowd is drawn, the firm's ability to recruit the population's most accurate members, and the fact that firms care about satisficing, not optimizing behavior. In so doing, the paper answers previous calls for a better understanding of how ideas are selected by crowd-based procedures (Afuah and Tucci 2012) and of how different decision-making structures lead to different performance outcomes (Gavetti et al. 2007). In practical terms, the analytical tools developed in this paper can be used to guide organization designers on how to profitably use crowds to select ideas.

It is well known that organizational performance depends on possessing foresight (Denrell and Fang 2010, Csaszar and Laureiro-Martinez 2018)—to know, for instance, which markets to enter, what resources to acquire, or what products to launch. However, foresight is a rare commodity, as uncertainty, complexity, and bounded rationality, all limit the predictive powers of individuals. The promise of organizations is that, by selecting an appropriate aggregation mechanism, they will be able to exhibit more foresight than can any individual. The aim of this study is to move research one step closer to fulfilling that promise.

References

Afuah, A., C. L. Tucci. 2012. Crowdsourcing as a solution to distant search. *Academy of Management Review* **37**(3) 355–375.

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

Baumann, O., D. Martignoni. 2011. Evaluating the new: The contingent value of a pro-innovation bias. *Schmalenbach Business Review* **63**(4) 393–415.

Baumann, O., N. Stieglitz. 2014. Rewarding value-creating ideas in organizations: The power of low-powered incentives. *Strategic Management Journal* **35**(3) 358–375.

Bayus, B. L. 2013. Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community. *Management Science* **59**(1) 226–244.

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.

Berend, D., J. Paroush. 1998. When is Condorcet's jury theorem valid? *Social Choice and Welfare* **15**(4) 481–488.

Berend, D., L. Sapir. 2005. Monotonicity in Condorcet Jury Theorem. *Social Choice and Welfare* **24**(1) 83–92.

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.

Black, D. 1958. *The Theory of Committees and Elections*. Cambridge University Press, Cambridge, UK.

Boland, P. J. 1989. Majority systems and the Condorcet jury theorem. *The Statistician* **38**(3) 181–189.

Bonabeau, E. 2009. Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review* **50**(2) 45–52.

Boudreau, K. J., N. Lacetera, K. R. Lakhani. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science* **57**(5) 843–863.

Chesbrough, H. W. 2003. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press, Boston, MA.

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

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* **24**(4) 1083–1101.

Csaszar, F. A., J. P. Eggers. 2013. Organizational decision making: An information aggregation view. *Management Science* **59**(10) 2257–2277.

Csaszar, F. A., D. Laureiro-Martinez. 2018. Individual and organizational antecedents of strategic foresight—a representational approach. Forthcoming in *Strategy Science*.

Denrell, J., C. Fang. 2010. Predicting the next big thing: Success as a signal of poor judgment. *Management Science* **56**(10) 1653–1667.

Dougherty, K. L., J. Edward. 2009. Odd or even: Assembly size and majority rule. *Journal of Politics* **71**(2) 733–747.

Erat, S., V. Krishnan. 2012. Managing delegated search over design spaces. *Management Science* **58**(3) 606–623.

Feddersen, T., W. Pesendorfer. 1997. Voting behavior and information aggregation in elections with private information. *Econometrica* **65**(5) 1029–1058.

Fernández, M., S. Williams. 2010. Closed-form expression for the Poisson-binomial probability density function. *IEEE Transactions on Aerospace and Electronic Systems* **46**(2) 803–817.

Galton, F. 1907. Vox populi. *Nature* **75**(1949) 450–451.

Gast, A., M. Zanini. 2012. The social side of strategy. *McKinsey Quarterly* (2) 82–93.

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.

Girotra, K., C. Terwiesch, K. T. Ulrich. 2010. Idea generation and the quality of the best idea. *Management Science* **56**(4) 591–605.

Goldstein, D. G., R. P. McAfee, S. Suri. 2014. The wisdom of smaller, smarter crowds. *Proceedings of the Fifteenth ACM Conference on Economics and Computation*. ACM, New York, 471–488.

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

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

Hong, Y. 2013a. On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis* 41–51.

Hong, Y. 2013b. *poibin: The Poisson Binomial Distribution*. URL http://cran.r-project.org/package=poibin. R package version 1.2.

Howe, J. 2008. *Crowdsourcing*. Crown Business, New York.

Jeppesen, L. B., K. R. Lakhani. 2010. Marginality and problem-solving effectiveness in broadcast search. *Organization Science* **21**(5) 1016–1033.

Karotkin, D., J. Paroush. 2003. Optimum committee size: Quality-versus-quantity dilemma. *Social Choice and Welfare* **20**(3) 429–441.

King, A., K. R. Lakhani. 2013. Using open innovation to identify the best ideas. *MIT Sloan Management Review* **55**(1) 41–48.

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

Lakhani, K. R., Z. Kanji. 2008. *Threadless: The Business of Community*. Harvard Business School Case No. 608–707, Cambridge, MA.

Laughlin, P. R., A. L. Ellis. 1986. Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology* **22**(3) 177–189.

Marschak, J., R. Radner. 1972. *Economic Theory of Teams*. Yale University Press, New Haven, CT.

May, K. O. 1952. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica* **20**(4) 680–684.

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

Nitzan, S., J. Paroush. 1982. Optimal decision rules in uncertain dichotomous choice situations. *International Economic Review* **23**(2) 289–297.

Peter, L. J., R. Hull. 1969. *The Peter Principle*. W. Morrow, New York.

Reitzig, M., B. Maciejovsky. 2015. Corporate hierarchy and vertical information flow inside the firm—a behavioral view. *Strategic Management Journal* .

Reitzig, M., O. Sorenson. 2013. Biases in the selection stage of bottom-up strategy formulation. *Strategic Management Journal* **34**(7) 782–799.

Sah, R. K., J. E. Stiglitz. 1988. Committees, hierarchies and polyarchies. *Economic Journal* **98**(391) 451–470.

Schwenk, C. R. 1988. The cognitive perspective on strategic decision-making. *Journal of Management Studies* **25**(1) 41–55.

Shapley, L., B. Grofman. 1984. Optimizing group judgmental accuracy in the presence of interdependencies. *Public Choice* **43**(3) 329–343.

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

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

Terwiesch, C., K. T. Ulrich. 2009. *Innovation Tournaments: Creating and Selecting Exceptional Opportunities*. Harvard Business Press, Boston, MA.

Tetlock, P. E. 2005. *Expert Political Judgment*. Princeton University Press, Princeton, NJ.

Thompson, D. N. 2012. *Oracles: How Prediction Markets Turn Employees Into Visionaries*. Harvard Business Review Press, Boston, MA.

Tischler, L. 2002. He struck gold on the net (really). *Fast Company* **59** 40–49.