Diversification and Information in Contests

Jorge Lemus* and Emil Temnyalov†

September 2, 2018

Abstract

We introduce a novel model of variety in contests, where agents can pursue different technologies to compete, but there is uncertainty regarding which technology will be implemented ex post and succeed in the contest. We study optimal information disclosure and diversification in the presence of such technological uncertainty. The principal can credibly reveal some information about the technologies, affecting the agents' choices. We completely characterize the optimal information disclosure policy, as a function of the set of possible information structures that the principal can implement and the agents' prior beliefs about the technologies. The optimal policy can be maximally or partially revealing, or completely uninformative, depending on: (i) the option value of diversification; (ii) the quality of the principal's information; and (iii) the extent of technological uncertainty. Our results can be applied to a variety of settings such as innovation contests, labor tournaments and procurement, where competing agents are uncertain about different possible ideas, projects, or the preferences of the principal.

Keywords: contests, variety, diversification, information disclosure

*University of Illinois Urbana-Champaign, Department of Economics. jalemus@illinois.edu

†University Technology Sydney, Business School. emil.temnyalov@uts.edu.au

1 Introduction

Contests describe a wide variety of different economic settings including innovation, procurement, promotions within organizations, and lobbying. In practice uncertainty plays a prominent role in many of these settings. For instance, in an innovation context, innovators face technological uncertainty and do not know with certainty which technologies might succeed or fail; in procurement, a firm may face uncertain preferences or objectives of the procurer (e.g. a defense procurement firm may not fully know the preferences of the procurer over different possible characteristics of the product); in organizations, workers may face uncertainty regarding on the impact of different projects on getting a promotion. In such settings, the principal may be able to reveal information to the agents that affects their beliefs regarding the different alternatives, and hence agents' may change their actions in a way that improves revenue or efficiency.

To study such settings we consider a contest with multiple possible "technologies," i.e. different approaches to solve a problem, methods, projects, or ideas, where only one of these technologies will be implemented ex post by the principal. We study optimal disclosure in a Bayesian persuasion-type framework, where the principal can resolve some of this uncertainty by credibly revealing information to the agents regarding the different technologies they might pursue in the contest. Revealing more information may induce an equilibrium where agents focus their efforts more efficiently (e.g., on a more promising technology), but it may induce too little diversity of technologies pursued in equilibrium, relative to what would be optimal for the principal. This creates a trade-off for the principal when revealing information, as diversity may be desirable in an environment where technologies are uncertain, since the principal may want to "hedge" her bets.

Consider for example the Netflix Prize, where Netflix ran a contest to improve the algorithm it uses to give recommendations to its users. In designing this contest, Netflix revealed a subset of all its available data on users' movie ratings, to allow contestants to test their own prediction methods and evaluate how promising different possible approaches would be. From Netflix's perspective there is clearly some option value to procuring a number of different algorithms, in addition to whichever algorithm performs best given their current data, as some algorithms may turn out to be more valuable in the future, when Netflix has more data on users' preferences. Thus an interesting contest design question for Netflix is how much of their available data to reveal to the competitors. Revealing more may allow contestants to make better predictions based on Netflix's current data, but it may also induce too little diversification across different methods, if their current data clearly favors one method over another. Indeed, when running the contest Netflix only offered a small sample of their available data.

Websites such as Kaggle and DrivenData now offer all firms the ability to run very similar prediction contests and to "outsource" their data analytics. These contests also feature technological uncertainty and information disclosure, as contestants receive information regarding the performance of different possible approaches (e.g. machine learning, regression methods, or other prediction algorithms), and they can exchange this information (usually through public forums). Similarly, in the Netflix competition the best approach in a preliminary stage was one proposed by a user with the pseudonym "Simon Funk," which used a "singular value decomposition" prediction method. Once the performance of this approach became known, "Everybody adopted Simon's ideas..." said one of the contest participants.[1]

A similar example is the Hyperloop Pod Competition, sponsored by Space X, with the goal of testing prototypes for the Hyperloop.[2] The competition was run in several stages, with the results made publicly available. A wide variety of technologies were explored, including pod designs that use air bearings, others that use magnetic levitation, and others that use high speed wheels. The technology used by the winner of the first part of the competition was a pod design that used an electrodynamic suspension system to levitate and an axial compressor to minimize aerodynamic drag. Is it optimal for the contest designer to disclose this information after the first round? One possible drawback is that other teams may be tempted to abandon their approaches and switch to the currently most promising approach, as a reaction to good news about one particular technology. However, given that there is uncertainty regarding which technology will turn out to be the best in the long run, the current best technology may not be implementable ex post, so the contest designer may see value in a variety of approaches.

This paper studies how technological uncertainty affects the agents' choices of technologies in a contest, and whether the principal can benefit from information disclosure that is designed to change the agents' beliefs about the alternatives. We assume that

[1] https://www.thrillist.com/entertainment/nation/the-netflix-prize

[2] https://en.wikipedia.org/wiki/Hyperloop_pod_competition

the principal can design a public experiment regarding the different technologies, which will reveal information about the likelihood that they are successful or valuable. The principal has the ability to commit to such a signal structure ex ante, and cannot hide information ex post. We find that under very general conditions relating to the features of the environment, the principal strictly benefits from the optimal information design rule.

We precisely characterize the optimal information structure that maximizes the principal's expected payoff from the contest. The main trade-off for the principal is one between diversification and focus: revealing more precise information induces agents to focus on more promising technologies, but if the agents optimally over-react to such information in equilibrium, then this may lead to too little diversification from the principal's perspective. As a result of this trade-off we find that the optimal policy can be either maximally informative, partially informative or completely uninformative, depending on different features of the environment. Whether the principal wants to reveal or hide information depends critically on three factors: (i) the value of technological diversity; (ii) the quality of the principal's information; and (iii) the extent of technological uncertainty.

First, the value of technological diversity reflects the fact that the principal herself does not know which approach will be implemented. Hence developing multiple different approaches has some option value: even if one is more promising than another ex ante, the latter may turn out to be more valuable in the long run. The larger this value of diversification is, the more likely it is that the principal chooses to hide information.

Second, the quality of the principal's information matters: if she can design a very informative experiment, which will reveal with very high probability which technology will be implemented ex post, then she is more likely to want to reveal such information. In practice, however, the principal may not have access to such very informative signals, in which case she may prefer to hide information.

Third, the extent of technological uncertainty reflects how similar or asymmetric the different approaches are a priori. If the agents' beliefs about the technologies are ex ante very asymmetric, the principal may want to reveal information to either reinforce or weaken the extent of this asymmetry. The more symmetric the technologies are ex ante, the more likely it is that the principal will choose to hide information.

2 Related Literature

This paper relates most closely to a recent literature on diversification or variety in contests. Letina and Schmutzler (2017) study a Hotelling-style contest where 2 contestants can choose what technologies to develop, with the Hotelling space representing different alternatives over which the principal has uncertain preferences. They characterize the optimal prize structure when the designer wants to induce a variety of approaches. Our paper offers a new model of variety, which does not rely on the Hotelling framework and provides a more tractable way to model variety across many technologies and with many players. Our results are complementary to theirs, in that we focus on the problem of information design, which can also be used to induce variety, rather than on the optimal prize structure.

A related literature studies variety in other settings. Terwiesch and Xu (2008) incorporate diversity into the preferences of the contest designer and show that more participants may be preferred. Boudreau et al. (2011) empirically test the effect of the number of participants on diversity. Letina (2016) studies the effect of market competition and mergers on variety, and finds conditions such that the research portfolio under market competition features too much (or too little) variety. Toh and Kim (2013) study how aggregate uncertainty affects technological diversification within a firm. They find that a firm's technology becomes more specialized under greater uncertainty. Related to this, Krishnan and Bhattacharya (2002) study how a firm should design a product when there are several uncertain alternatives for the product's underlying technology.

Because we focus on information design, our paper also contributes to the literature on disclosure and feedback in contests. This literature focuses on information regarding the agents' characteristics or actions, rather than information about the underlying technologies, as in our paper. For instance, Aoyagi (2010) studies a dynamic tournament and compares the effort provision by agents under full disclosure of their performance (i.e., players observe their relative position) and no information disclosure. Ederer (2010) adds private information to this setting, and Klein and Schmutzler (2016) add other decisions regarding the allocation of prizes and alternative performance evaluations. Fu et al. (2016) and Xin and Lu (2016) study optimal information disclosure regarding agents' entry decisions in contests. Zhang and Zhou (2016) study information disclosure regarding one player's effort costs. Kovenock et al. (2015) studies the effect of players sharing information throughout the contest. Feedback in dynamic contests has been recently studied by Bimpikis et al. (2014), and Benkert and Letina (2016). Recent empirical work assessing the effect of performance feedback on competition outcomes includes Gross (2015), Gross (2017), Huang et al. (2014), Kireyev (2016), Bockstedt et al. (2016), and Lemus and Marshall (2017).

This paper also relates to R&D models with multiple risky technologies. Dasgupta and Maskin (1987) shows that in a winner-takes-all competition, the equilibrium allocation of research on correlated projects is too high relative to the socially efficient allocation, so there is less diversification in equilibrium. Bhattacharya and Mookherjee (1986) present a similar framework, but they study the level of risk taken by the firms finding that the optimal research strategy may feature excessive or insufficient risk taking, depending on the level of risk aversion and the shape of the distribution of research outcomes. Cabral (1994) shows that when the competition is not winner-takes-all, the level of risk taking is lower than the socially optimal level. Cabral (2003) explore the same question in a dynamic environment, showing that a follower firm takes more risk than the leader. Our paper contributes to the literature by studying information disclosure by the contest designer, as in Kamenica and Gentzkow (2011).

A recent literature has explored information design in games more generally. This work includes Zhang and Zhou (2016), Mathevet et al. (2017), Laclau and Renou (2016), Alonso and Câmara (2016) among others, and also models of contests for experimentation, such as Halac et al. (2017).

Our paper focuses on a single element of contest design, information disclosure, which relates broadly to the literature of contests design, where the goal is to study the effect of alternative designs on players' incentives. These literature includes the work of Taylor (1995) and Fullerton and McAfee (1999) on restricting the number of competitors in winner-takes-all tournaments, Moldovanu and Sela (2001) on the optimal number of prizes, Che and Gale (2003) on both number of prizes and the number of participants.

3 Model

A continuum of agents, $i \in [0, 1]$, compete in a contest that awards a single prize V. Each agent has one indivisible unit of effort and chooses to allocate it to one of N

different technologies, $j \in N$. Each technology represents a different "technological approach" and only one of these approaches will be implemented ex post. Agents hold a prior belief $p_j \in [0,1]$ that technology j is the correct approach, with $\sum_j p_j = 1$. The probabilities $\{p_j\}_j$ are commonly known to all agents. We model competition among agents in a reduced form. Given that x_j agents allocate their effort towards technology j, the probability that one of these agents wins the contest is the same. In other words, we assume the contest is anonymous: if two agents both choose a successful technology, they have the same chances of winning the contest. Hence the probability of an individual winning the contest is uniform over the agents who choose the technology that is implemented ex post.

Assumption 1. *If a measure of x agents work on the implemented technology, each of them has a probability (density) $s(x) = \frac{1}{x}$ of winning the contest.*

The principal's payoff depends on how many agents allocate their effort to the ex post successful technology. If x_j agents choose to work on technology j, and technology j turns out to be the correct approach that is implemented ex post, the principal's payoff ex post is

$$v(x) = f(x_j),$$

where $f(\cdot)$ is an increasing and concave function, which represents the gains from agents' efforts or investments into the chosen technology. This payoff structure is a reduced form representation of a model where agents take draws out of some distributions, and $f(x_j)$ represents the best draw when a mass x_j of agents draw from a distribution that corresponds to a successful technology. An ex post unsuccessful technology (i.e., not feasible or not valuable) results in unproductive or wasteful effort, from the designer's perspective.

With any beliefs $\{q_j\}_j$ regarding the technologies $j \in N$, the first-best allocation of agents solves the following problem:

$$x^{FB} \in \arg\max_{x \in [0,1]^N} \sum_j q_j f(x_j) \quad \text{s.t.} \quad \sum_j x_j = 1. \tag{1}$$

Proposition 1 (First best allocation)**.** *When $f(\cdot)$ is increasing, differentiable and con-*

cave, the solution to Problem (1) is characterized by

$$\frac{f'(\bar{x}_i)}{f'(\bar{x}_j)} = \frac{q_j}{q_i} \quad \forall i, j \in N.$$

Proposition 1 shows that the first-best solution for the principal equates the ratios of marginal gains from each technology to the inverse ratios of their probabilities of implementation. This is intuitive, since the first best is to allocate efforts in a way that equalizes the marginal expected gains from the technologies. Given that f is increasing, the contest designer allocates more agents to technologies that are more promising; and given that F is concave, the contest designer diversifies and allocates agents across multiple technologies, since they are uncertain.

The equilibrium allocation of agents, given any common beliefs $\Theta = (\theta_1, ..., \theta_N)$ about the technologies, is characterized in the next lemma.

Proposition 2 (Equilibrium characterization)**.** *If agents hold beliefs Θ about the technologies $j \in N$, the equilibrium mass of agents working on technology j is*

$$x_j = \theta_j.$$

Proof. Consider an allocation of agents where x_j agents work on technology j, with $\sum_{j=1}^{N} x_j = 1$. For this allocation to be an equilibrium, no agent must have incentives to deviate, so each technology must give each agent the same payoff which implies

$$\frac{\theta_j}{x_j} = \frac{\theta_k}{x_k}, \quad \text{for all } j, k \in \{1, ..., N\}.$$

This condition is equivalent to $x_j = \frac{\theta_j}{\theta_k} x_k, \quad$ for all $j = 1, ..., N$, so adding over j and using that $\sum_{j=1}^{N} x_j = \sum_{j=1}^{N} \theta_j = 1$, we obtain $x_k = \theta_k$. Since k was arbitrary, the result follows. □

Proposition 2 shows that in equilibrium more agents are allocated to the most promising technology, but competition among agents pushes some of them to work on less promising technologies: the chances of winning are larger when fewer agents work on one technology. Notice also that the equilibrium is generally inefficient.[3] That means

[3] Except in the special case where $f(x) = \log(x)$, where $\frac{f'(\bar{x}_i)}{f'(\bar{x}_j)} = \frac{\bar{x}_j}{\bar{x}_i} = \frac{\theta_j}{\theta_i} = \frac{x_j}{x_i}$.

the equilibrium allocation, relative to the first-best allocation, features too much or too little diversification. For instance, when $f(x) = x^a$ for some $a \in (0,1)$, from Proposition 1 and Proposition 2 we have:

$$\left(\frac{\bar{x}_j}{\bar{x}_i}\right)^{1-a} = \frac{\theta_j}{\theta_i} = \frac{x_j}{x_i}.$$

Suppose that technology 1 is the less promising one, i.e. $\theta_1 < \theta_j$ for $j = 2, ..., N$. In the first-best allocation, $\bar{x}_j = \left(\frac{\theta_j}{\theta_1}\right)^{\frac{1}{1-a}} \bar{x}_1$, and in equilibrium $x_j = \frac{\theta_j}{\theta_1} x_1$. Given that $a \in (0,1)$ this implies $\bar{x}_j = \gamma x_j$, where $\gamma = \left(\frac{\theta_j}{\theta_1}\right)^{\frac{a}{1-a}} > 1$ because $\theta_j > \theta_1$. Hence, $\bar{x}_j > x_j$, so in equilibrium agents "under-react" to good news about technology $j > 1$ (and too many agents work on the least likely technology). The reason is that competition is "too fierce" for the agents relative to the marginal gain of allocating an extra agent to technology $j > 1$ for the contest designer. Similarly, when $f(x) = -(1 - x^{1/a})^a$ for some $a > 1$, agents "over-react" to asymmetries in the technologies.

Given that competing agents' incentives and the preference of principal are misaligned, we explore whether the principal can design an experiment to persuade agents to work towards different technologies.

4 Information design

The above analysis holds for any beliefs of the principal and the agents. We now turn to the analysis of optimal information disclosure, when the designer can credibly reveal to the agents information about the feasibility or value of the technologies, before they choose how to allocate their efforts.

In a situation where technology-specific prizes are not contractible, can the contest designer disclose information to improve the equilibrium allocation of agents across technologies? We study this question in a Bayesian persuasion-style information design framework: suppose the principal can design a public experiment which will reveal information about the technologies to all agents. We assume ex ante commitment to an information disclosure policy, like in the special case of Bayesian persuasion.

In particular, an experiment is a signal structure $s = (M, \widetilde{G}(m|j))$, where M is an arbitrary set of messages and $\widetilde{G}(m|j)$ is a distribution over the messages, in the state that technology j is implemented ex post. Let S be the set of all such signal structures available to the principal. Note that we do not assume that the principal has access to every signal structure (for example, a perfectly informative signal might be impossible, which is the case in many practical applications). Instead, we solve for the optimal information disclosure policy for *any* set of available signals. Moreover, the set S allows for discrete or continuous posteriors, and it allows for both partitional or noisy signal structures. Each signal structure s induces some ex ante distribution over posteriors, $G_s(\Theta)$, and we denote the set of posteriors in the support of that signal as $\mathcal{P}_s$, with generic elements $\Theta \in \mathcal{P}_s$. Let $\mathcal{P}_S \equiv \cup_{s\in S}\mathcal{P}_s$ denote the set of all posteriors that can be induced by some signal.

The value of information disclosure can be analyzed in terms of the beliefs that any disclosure induces for the agents and the designer's payoff as a function of these beliefs. Recall that in equilibrium, $x_j = \theta_j$, so the contest designer's expected payoff from posterior Θ is

$$\nu(\Theta) \equiv \sum_j \theta_j f(\theta_j). \tag{2}$$

The value of information disclosure is described by the concavity/convexity of $\nu(\Theta)$. Let $\hat{\nu}(\Theta, \mathcal{P}_S)$ be the concave closure of $\nu(\Theta)$ over $\mathcal{P}_S$. Then the principal strictly benefits from *some* information disclosure whenever $\hat{\nu}(\theta_0, \mathcal{P}_S) > \nu(\theta_0)$ around the prior θ_0. We can determine whether this value function is concave or convex over certain regions by looking at whether its Hessian is negative or positive semidefinite.

Lemma 1. *Define $g(\theta) = \theta f(\theta)$. Then, $\nu(\Theta)$ is concave at $\Theta = (\theta_1, ..., \theta_N)$ if and only if $g''(\theta_j) < 0$ for all $j = 1, ..., N$.*

Proof. Let $H(\Theta)$ be the matrix of second derivatives of $\nu(\Theta)$. Given the separability of the function $\nu(\cdot)$ the matrix $H(\Theta)$ is diagonal, with $\frac{\partial^2 \nu(\Theta)}{\partial^2 \theta_j}$ in the j-th row and column. Concavity can be verified by checking that $z^T H(\Theta) z < 0$ for all $z \in R^N \setminus \{0\}$. This condition is equivalent to

$$\sum_{j=1}^{N} z_j^2 g''(\theta_j) < 0.$$

Thus, a necessary and sufficient condition for this to hold is $g''(\theta_j) < 0$ for all $j = 1, ..., N$. $\square$

Lemma 1 implies that we need to study the concavity of $\theta f(\theta)$ to determine whether or not there are gains from information disclosure. This turns out very easy to characterize:

$$\frac{\partial^2[\theta f(\theta)]}{\partial\theta^2} < 0 \Leftrightarrow 2f'(\theta) + \theta f''(\theta) < 0 \Leftrightarrow 2 < \frac{-\theta f''(\theta)}{f'(\theta)}.$$

Hence whether $\nu(\Theta)$ is locally concave at Θ depends on the Arrow-Pratt relative risk aversion coefficient associated to $f(\cdot)$, around the belief θ_j, defined by

$$r(x) \equiv \frac{-xf''(x)}{f'(x)}.$$

This coefficient captures the value of diversity to the principal: when $r(x)$ is large, diversifying across different technologies is more valuable, because the returns to effort on each technology diminish quickly.

This can be illustrated with some simple examples. It is easy to see from Lemma 1 that for $f(x) = x^a$, with $a \in (0,1)$ the function $\nu(\cdot)$ is globally convex, so there are always gains from information disclosure, for any prior. On the other hand, for $f(x) = -(1-x^{1/a})^a$, for $a > 1$, the function $\nu(\cdot)$ can be concave around the middle and convex near the extremes—for a large enough, $f(\cdot)$ is close to linear around the extremes, so $\nu(\Theta)$ is convex there, whereas $f(\cdot)$ is very concave around the middle, so $\nu(\Theta)$ is concave there.

Intuitively, whether there are gains from information disclosure or not (i.e. whether $\nu(\cdot)$ is locally convex or concave) depends on the value of diversification across technologies to the principal. When $\nu(\cdot)$ is very concave, diversification is more valuable to the principal than to the agents, because the equilibrium allocation of effort, for any given belief Θ, is too responsive to differences among the technologies, relative to the principal's first-best allocation of effort across technologies. In this case agents "over-react" to differences among the technologies. Specifically, equilibrium efforts are characterized by $\frac{x_i}{x_j} = \frac{\theta_i}{\theta_j}$, whereas the first-best allocation is characterized by $\frac{f'(x_j^{FB})}{f'(x_i^{FB})} = \frac{\theta_i}{\theta_j}$, which for f concave enough implies that $\frac{x_i}{x_j}$ is farther towards the extremes than $\frac{x_i^{FB}}{x_j^{FB}}$. If the principal were to reveal information that induces a belief that some technologies are better than others, the equilibrium choices of agents would be too concentrated on more promising technologies, relative to the first-best allocation, and hence the principal may prefer not to reveal information.

In contrast, when $r(x)$ is relatively small, i.e. $f(\cdot)$ is relatively less concave, then the value of diversification to the principal is smaller, and so revealing information that induces more extreme beliefs is more valuable, and produces an allocation closer to the first-best. In this case agents "under-react" to differences among the technologies. Therefore the principal may prefer to reveal information that induces more extreme beliefs (as in the example with $f(x) = x^a$).

Our main result on information disclosure captures this intuition. Before we can state the result, we first prove some preliminary results and we define some additional notation which will then allow us to describe the optimal signal structure more simply.

Lemma 2. *For any set of signal structures S, the set of posteriors that can be induced by the principal, $\mathcal{P}_S = \cup_{s \in S} \mathcal{P}_s$, is convex.*

Proof. Consider any two posteriors $\Theta', \Theta'' \in \mathcal{P}_S$ induced by some messages m' and m'', from (possibly different) signal structures s' and s'', respectively. For any $\alpha \in (0, 1)$, the posterior $\alpha\Theta' + (1-\alpha)\Theta''$ can be induced with a signal structure s^* that sends a message m^* with probability α whenever s' would send m', and sends m^* with probability $1-\alpha$ whenever s'' would send m'', and sends any other arbitrary messages otherwise.

Conditional on observing a message m^*, each agent believes that with probability α the conditional probability of state j is Θ', and with probability $1-\alpha$ the conditional probability of state j is Θ''. Hence the agent's posterior is $\alpha\Theta' + (1-\alpha)\Theta''$, so the set $\mathcal{P}_S$ is convex. □

Intuitively, for any two posteriors $\Theta', \Theta'' \in \mathcal{P}_S$ that can be induced with some signal structures s' and s'', the principal can also induce any convex combination $\alpha\Theta' + (1 - \alpha)\Theta''$ with a signal structure that mixes s' and s'' in the right way. Therefore the set of feasible posteriors is a convex subset of the $N-1$ simplex.

Lemma 3. *The function $\nu(\Theta)$ has the following properties:*

(i) it has global maxima at the vertices of the $N-1$ simplex;

(ii) $\nu(\Theta)$ has at most one local maximum, at the center of the simplex, if $2f'(\theta_j) + \theta_j f''(\theta_j)$ crosses 0 at most once for all $\theta_j \in [0, \frac{1}{N}]$.

Proof. For part (i), note that at each vertex $\theta_i = 1$ for some $i \in N$ and $\theta_j = 0 \, \forall j \neq i$. Hence the values at the vertices are $\nu(0, ..., 0, 1, 0, ..., 0) = 1 \cdot f(1)$. Since $f(\cdot)$ is increasing, this is the largest possible value that the objective may obtain. Hence $\nu(\Theta)$ has global maxima at each vertex.

For part (ii), we will first show that $\nu(\Theta)$ has a critical point at the center of the simplex; we then discuss how whether that critical point is an interior local maximum or minimum depends on $2f'(\theta_j) + \theta_j f''(\theta_j)$, and how this expression determines the uniqueness of an interior maximum.

Suppose $\nu(\Theta) = \sum_{j=1}^{N-1} \theta_j f(\theta_j) - (1 - \sum_{j=1}^{N-1} \theta_j) f(1 - \sum_{j=1}^{N-1} \theta_j)$ has a local maximum in the interior of the simplex. This is characterized by the FOC with respect to θ_j:

$$f(\theta_j) + \theta_j f'(\theta_j) = f(1 - \sum_{j=1}^{N-1} \theta_j) + (1 - \sum_{j=1}^{N-1} \theta_j) f'(1 - \sum_{j=1}^{N-1} \theta_j).$$

Hence we have

$$f(\theta_j) + \theta_j f'(\theta_j) = f(\theta_i) + \theta_i f'(\theta_i)$$

for all $i, j \in N$, and a symmetric system of N nonlinear equations in N unknowns. It is clear that one solution to this system is given by $\theta_j = \frac{1}{N}$ for all $j \in N$. Hence $\nu(\Theta)$ has a critical point at the center of the simplex.

If the above system of equations has a unique solution, then that critical point is a local minimum, and $\nu(\Theta)$ must be convex everywhere, since it has global maxima at the vertices. On the other hand, if the system of equations has multiple solutions, some of them may be interior local maxima, with $\frac{\partial^2 \nu(\Theta')}{\partial \theta_j^2} = 2f'(\theta_j) + \theta_j f''(\theta_j) < 0$, since the objective $\nu(\Theta)$ is separable and its Hessian is a diagonal matrix.

Next, note that the objective $\nu(\Theta)$ has global maxima at the vertices, so it is locally convex near the vertex, i.e. $\frac{\partial^2 \nu(\Theta')}{\partial \theta_j^2} > 0$ in a neighborhood around each vertex. Therefore if $2f'(\theta_j) + \theta_j f''(\theta_j)$ crosses 0 at most once over $\theta_j \in [0, \frac{1}{N}]$, then all critical points to the left of some threshold $\bar{\theta}_j \in (0, \frac{1}{N}]$ with $2f'(\bar{\theta}_j) + \bar{\theta}_j f''(\bar{\theta}_j) = 0$ are minima, and any critical point to the right of $\bar{\theta}_j$ is an interior maximum. Moreover, because of single crossing in this case we must have a unique local maximum to the right of the threshold $\bar{\theta}_j$, if there exist any interior maxima. Thus the critical point at the center must be a local maximum if $2f'(\theta_j) + \theta_j f''(\theta_j)$ crosses 0.

Hence we have shown that if $\frac{\partial^2 \nu(\Theta')}{\partial \theta_j^2} = 2f'(\theta_j) + \theta_j f''(\theta_j)$ crosses 0 only once for all $\theta_j \in [0, \frac{1}{N}]$, then it does so from above, and there is a unique interior local maximum at the center, while any other critical points are local minima or saddle points.

□

Lemma 3 implies some useful properties of the principal's objective function.

First, it has global maxima with value $f(1)$ at the extreme points of the simplex, because at those extreme points the principal and agents know for certain which technology will be implemented, and agents in equilibrium allocate all effort to that technology. This immediately implies that if the agent has access to a perfectly informative signal, that signal would be optimal. However, such a signal need not be available to the principal in general—in many settings it is unrealistic to assume that the principal has a perfect signal, which can eliminate all uncertainty in the environment. In the Netflix Prize example, to implement such a signal Netflix would need to have infinite amounts of data on consumers' preferences. Our main result will characterize the optimal signal structure for any set S of possible signals, which need not include perfectly informative ones.

Second, the value function is generally convex towards the extremes, and may have a concave region around the center of the simplex, where it may have a local maximum. Whether such a concave region exists or not depends on the functional form of the production function, $f(x)$. The lemma provides a sufficient single-crossing condition to ensure that the objective has at most one interior local maximum, and does not oscillate: if $2f'(x) + xf''(x)$ crosses 0 at most once, then the objective either only has an interior local minimum, or it has a unique interior local maximum at the center. We can interpret this condition in terms of the relative risk coefficient, $r(x) \equiv \frac{-xf''(x)}{f'(x)}$: if $r(x)$ is well-behaved, then there is at most one interior local maximum. A sufficient condition for this would be to assume that the "relative prudence" coefficient,

$$rp(x) \equiv \frac{-xf'''(x)}{f''(x)}$$

satisfies $rp(x) < 3$ for all $x \in [0, \frac{1}{N}]$.

We maintain the assumption in Lemma 3 for the remainder of the analysis, noting that the characterization of an optimal signal structure does not critically rely on this assumption—one could also state a general characterization result without it, but the assumption allows us to more explicitly state the main result and the optimal distributions over posteriors.

Next, let

$$\underline{\theta}_i \equiv \inf\{\theta_i : (\theta_i, \theta_{-i}) \in \mathcal{P}_S\}$$
$$\overline{\theta}_i \equiv \sup\{\theta_i : (\theta_i, \theta_{-i}) \in \mathcal{P}_S\}$$

denote the smallest and largest feasible beliefs for each technology i. For convenience of exposition we will assume the set of possible posteriors $\mathcal{P}_S$ is symmetric (though individual signals need not be symmetric), hence $\underline{\theta}_i = \underline{\theta}_j = 1 - \overline{\theta}_i = 1 - \overline{\theta}_j$ for all i, j. This makes it easier to state the conditions for our main result, but is not generally necessary for the results.

Next, let

$$\widetilde{v} \equiv \sup\{\nu(\Theta) : \nu''(\Theta) < 0\}$$

denote the largest value of the value function over the region where it is concave.[4]

Finally, let

$$\hat{\Theta} \equiv \partial\{\Theta : \hat{\nu}(\Theta, \mathcal{P}_S) = \nu(\Theta)\}$$

denote the boundary of the set of all posteriors where the value function ν agrees with its concave closure $\hat{\nu}$ over $\mathcal{P}_S$.[5]

We can now characterize the optimal information design policy.

Proposition 3. *The optimal disclosure policy s^* is*

1. ***maximally informative*** *if $\nu(\overline{\theta}_i) \geq \widetilde{v}$ for all i; s^* induces a distribution with support consisting only of extreme points, with distribution G_s s.t. $\mathrm{E}_{G_s}[\Theta] = p$.*
2. ***partially informative*** *if $\nu(\overline{\theta}_i) < \widetilde{v}$ for all i and $\nu(p) \neq \hat{\nu}(p, \mathcal{P}_S)$; s^* induces a distribution with support consisting only of boundary points in $\hat{\Theta}$, with distribution G_s s.t. $\mathrm{E}_{G_s}[\Theta] = p$.*

[4] If this region is empty, $\widetilde{v} = -\infty$.

[5] Note that any point Θ such that $\hat{\nu}(\Theta, \mathcal{P}_S) \neq \nu(\Theta)$ must be a convex combination of points from $\hat{\Theta}$, by the definition of this boundary set.

3. ***completely uninformative*** *if* $\nu(\overline{\theta}_i) < \widetilde{v}$ *for all* i *and* $\nu(p) = \hat{\nu}(p, \mathcal{P}_S)$*;* s^* *induces the same posterior as the prior,* p*.*

Proof. We characterize the optimal signal in 3 distinct cases.

Case 1: $\nu(\overline{\theta}_i) \geq \widetilde{v}$ **for all** i**.** First, note that if $\nu(\overline{\theta}_i) \geq \widetilde{v}$ for all i, then the global maxima of $\nu(\Theta)$ over $\mathcal{P}_S$ are attained at the points where $\theta_i = \overline{\theta}_i$ for each $i \in N$. So the concave closure of $\nu(\Theta)$ over the region $\mathcal{P}_S$ is the plane that connects the vertices of $\mathcal{P}_S$.

Note that the vertex for technology i is given by the points where $\theta_i = \overline{\theta}_i$ and $\theta_j = \frac{1-\overline{\theta}_i}{N-1}$ for all $j \neq i$. Denote the vertex for technology i by $\hat{\Theta}_i = (\frac{1-\overline{\theta}_i}{N-1}, ..., \frac{1-\overline{\theta}_i}{N-1}, \overline{\theta}_i, \frac{1-\overline{\theta}_i}{N-1}, ..., \frac{1-\overline{\theta}_i}{N-1})$.

Then the optimal signal s^* only induces posteriors $\{\hat{\Theta}_i\}_{i \in N}$. For any arbitrary prior $p \in \mathcal{P}_S$, Bayesian consistency of the posteriors then uniquely pins down the distribution over $\{\hat{\Theta}_i\}_{i \in N}$. In particular, there exists a unique distribution G_{s^*} with support $\{\hat{\Theta}_i\}_{i \in N}$ such that $\mathrm{E}_{G_{s^*}}[\Theta] = p$. This characterizes the distribution of posteriors induced by the optimal signal s^* in this case.

Case 2: $\nu(\overline{\theta}_i) < \widetilde{v}$ **for all** i **and** $\nu(p) \neq \hat{\nu}(p, \mathcal{P}_S)$**.** First, note that $\nu(\overline{\theta}_i) < \widetilde{v}$ for all i, then this implies that there exists a region in the interior where $\nu\Theta$ is concave, or else $\widetilde{v} = -\infty$. Moreover, the global maximum of $\nu(\Theta)$ over $\mathcal{P}_S$ is attained at the center of the simplex, $(\frac{1}{N}, ..., \frac{1}{N})$, by Lemma 3.

Next, $\nu(p) \neq \hat{\nu}(p, \mathcal{P}_S)$ implies that $\nu(\Theta)$ has a convex region towards the vertices of $\mathcal{P}_S$, and $\nu(p) < \hat{\nu}(p, \mathcal{P}_S)$ around the prior p. So the concave closure of $\nu(\Theta)$ over the region $\mathcal{P}_S$, denoted $\hat{\nu}(\Theta, \mathcal{P}_S)$, connects points in the boundary $\hat{\Theta} = \partial\{\Theta : \hat{\nu}(\Theta, \mathcal{P}_S) = \nu(\Theta)\}$ and the values at the vertices, $\nu(\overline{\theta}_i)$.

Then the optimal signal s^*, for a given prior p, only induces posteriors from the boundary, $\hat{\Theta}$, including both interior points and vertices of $\mathcal{P}_S$. Bayesian consistency then determines the distribution over $\hat{\Theta}$, with G_{s^*} such that $\mathrm{E}_{G_{s^*}}[\Theta] = p$. Note that this distribution need not be unique, depending on the shape of the boundary $\hat{\Theta}$, but the value $\hat{\nu}(p, \mathcal{P}_S)$ is unique, i.e. all optimal distributions are payoff-equivalent, and all induce posteriors in $\hat{\Theta}$.

Case 2: $\nu(\overline{\theta}_i) < \widetilde{v}$ for all i and $\nu(p) = \hat{\nu}(p, \mathcal{P}_S)$. First, note that $\nu(\overline{\theta}_i) < \widetilde{v}$ for all i, then this implies that there exists a region in the interior where $\nu\Theta$ is concave, or else $\widetilde{v} = -\infty$. Moreover, the global maximum of $\nu(\Theta)$ over $\mathcal{P}_S$ is attained at the center of the simplex, $(\frac{1}{N}, ..., \frac{1}{N})$, by Lemma 3.

Next, $\nu(p) = \hat{\nu}(p, \mathcal{P}_S)$ implies that $\nu(\Theta)$ is concave around the prior p, and the concave closure of $\nu(\Theta)$ over the region $\mathcal{P}_S$, denoted $\hat{\nu}(\Theta, \mathcal{P}_S)$, is equal to $\nu(\Theta)$. Hence the optimal signal induces a single posterior, $\Theta = p$, with a degenerate point distribution.

□

The optimal signal structure is characterized with 3 different cases. These cases depend on three key features of the environment: (i) the value of diversification for the principal; (ii) how informative are the principal's signals; and (iii) the extent of technological uncertainty. We explain the intuition for each of these features in turn.

First, the value of diversity matters because it determines whether the value function has a concave region with a local maximum $\widetilde{v}$ in the center. That is, when diversification is very valuable, i.e. the production function $f(x)$ is very concave and the $r(x)$ coefficient is large enough, then the principal may benefit from not inducing extreme posteriors. Such extreme posteriors generally lead the agents to over-react to asymmetries in the technologies and concentrate too much on the most promising one, from the principal's perspective. A stronger principal preference for diversity corresponds to a larger $\widetilde{v}$, while a weaker preference for diversity corresponds to a lower (or $-\infty$) value of $\widetilde{v}$. Hence lower values for diversity make case 1 more likely to hold, where revealing maximally informative signals is optimal; larger values for diversity make the principal less likely to reveal more information that induces extreme posteriors. Overall, a preference for diversity leads to less information disclosure.

Second, the informativeness of S matters, because it determines how large $\nu(\overline{\theta}_i)$ is. Having access to more informative signals, which induce more extreme posteriors, makes information disclosure more valuable. Recall that $\nu(\Theta)$ has global maxima at the vertices of the simplex, so at the extreme, the principal would want to reveal a perfectly informative signal, if it could do so. Even if such a signal is not available, if the principal has sufficiently informative signals that are close to perfectly informative, then $\nu(\overline{\theta}_i)$ is relatively large, and hence case 1 is more likely to hold. Overall, access to more informative signals lead to more information disclosure.

Third, technological uncertainty matters, because whether the a priori similarity or asymmetry of the technologies determines whether the value of diversification is large or not. If it is the case that diversification is valuable and the principal's signals are not informative enough, i.e. $\nu(p_1) < \widetilde{v}$, then asymmetries in the technologies make information disclosure more valuable. That is, if technology 1 is much better than 2 a priori, then the prior is such that $p_1 \in (\hat{\Theta}, \overline{\theta}_1)$, and the principal benefits from disclosure. In particular, in this case the principal wants to either reveal that technology 1 is indeed much better than 2 (pushing the belief up to $\overline{\theta}_1$), or reveal that the technologies are more similar than a priori (pushing the belief down to $\hat{\Theta}$); the corresponding probabilities for these two possibilities depend on the asymmetry (with larger p_1 implying more probability on the posterior $\overline{\theta}_1$). On the other hand, if the technologies are symmetric enough, i.e. $p_1 \leq \hat{\Theta}$, then the optimal signal reveals no information at all, and the principal does not benefit from information disclosure, because the value of diversity is relatively large, and more extreme beliefs would lower the principal's payoff. Overall, a priori asymmetries of technologies lead to more information disclosure.

4.1 The special case with $N = 2$

We can nicely illustrate the main result with the special case where $N = 2$. With $N = 2$, the prior is some $p_1 \in (0, 1)$, and the set of possible posteriors is $\theta_1 \in [\underline{\theta}_1, \overline{\theta}_1]$. Every signal structure $s \in S$ induces some distribution over $\mathcal{P}_S = [\underline{\theta}_1, \overline{\theta}_1]$. Given a realization θ_1, agents then choose technologies according to the equilibrium characterization in Proposition 2, and the principal's expected payoff is given by Equation 2. With $N = 2$ and $p_1 \geq 0.5$, we have

$$\widetilde{v} = \sup\{\nu(\theta_1) : \nu''(\theta_1) < 0\}$$
$$p_1'' = \sup\{\theta_1 : \nu''(\theta_1) \leq 0\}$$
$$\hat{\Theta} = \hat{\theta}_1 = \sup\{\theta_1 \in [\underline{\theta}_1, \overline{\theta}_1] : \hat{\nu}(\theta_1, \mathcal{P}_S) = \nu(\theta_1)\}$$

Corollary 1. *Suppose $N = 2$ and w.l.o.g. the prior is $p_1 \geq 0.5$. The optimal disclosure policy s^* is*

1. ***maximally informative*** *if* $\nu(\overline{\theta}_1) \geq \widetilde{v}$*;* s^* *induces a binary posterior distribution* $(\underline{\theta}_1, q; \overline{\theta}_1, 1-q)$ *with* $q\underline{\theta}_1 + (1-q)\overline{\theta}_1 = p_1$*.*
2. ***partially informative*** *if* $\nu(\overline{\theta}_1) < \widetilde{v}$ *and* $p_1 \in (p_1'', \overline{\theta}_1)$*;* s^* *induces a binary posterior distribution* $(\hat{\theta}_1, q; \overline{\theta}_1, 1-q)$ *with* $q\hat{\theta}_1 + (1-q)\overline{\theta}_1 = p_1$*.*
3. ***completely uninformative*** *if* $\nu(\overline{\theta}_1) < \widetilde{v}$ *and* $p_1 \leq \hat{\theta}_1$*;* s^* *induces a degenerate posterior distribution,* $(p_1, 1)$*.*

To illustrate these cases, consider the example with $f(x) = -(1-x^a)^{\frac{1}{a}}$, which induces the value function plotted in Figure 1 below. Here the principal's value function is concave in some region around the middle, but convex towards the extremes.

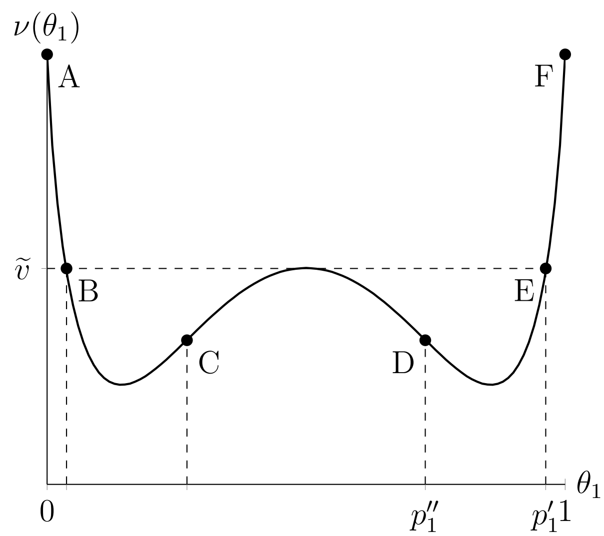

Figure 1: The value function $\nu(\theta_1)$ with $f(x) = -(1-x^a)^{\frac{1}{a}}$, with $a = 0.35$.

Here the structure of the principal's optimal signal depends on the quality or informativeness of the signals in S (i.e. on $\mathcal{P}_S$), on the value of diversification to the principal (i.e. on $\widetilde{v}$), and on the a priori asymmetry of the technologies (i.e. on the prior p_1).

First, if S includes signals that are informative enough, so that $\mathcal{P}_S$ includes posteriors close enough to 0 and 1, then we have $\nu(\underline{\theta}_1) > \widetilde{v}$ and $\nu(\overline{\theta}_1) > \widetilde{v}$. Graphically, $\underline{\theta}_1$ lies somewhere between points A and B, and $\overline{\theta}_1$ lies somewhere between points E and F in the figure. Here the concave closure of $\nu(\theta_1)$ is a line that connects $\nu(\underline{\theta}_1)$ and $\nu(\overline{\theta}_1)$. Then the optimal signal is one that reveals $\underline{\theta}_1$ with some probability q and $\overline{\theta}_1$ with the remaining probability $1-q$, where q is such that the expected posterior is equal to the prior, p_1. This is the first case in Corollary 1.

Similarly, if the value of diversification to the principal is relatively low, then $\widetilde{v}$ is low enough and we also have $\nu(\underline{\theta}_1) > \widetilde{v}$ and $\nu(\overline{\theta}_1) > \widetilde{v}$. Thus the optimal signal is also maximally informative, as in the first case of Corollary 1. The value of diversification to the principal is relatively small (i.e. the $r(x)$ coefficient is low enough), so revealing information to the agents increases the principal's expected value. Because of this, the agents' equilibrium allocation of efforts under-reacts to asymmetries in the technologies, so the principal benefits from inducing as extreme posteriors. This leads to an optimal disclosure rule that mixes between the 2 most extreme posteriors possible within $\mathcal{P}_S$. Graphically, if $f(x)$ is relatively less concave and the $r(x)$ coefficient is low enough, then the concave region around the center is relatively low, which pushes points B and E closer towards the middle.

Second, suppose the available signals are not as informative, and the set of feasible posteriors $\mathcal{P}_S$ is narrow enough that $\nu(\underline{\theta}_1) < \widetilde{v}$ and $\nu(\overline{\theta}_1) < \widetilde{v}$. Moreover, suppose the technologies are asymmetric enough, so that $p_1 > p_1''$. Graphically, $\underline{\theta}_1$ lies somewhere between points B and C, and $\overline{\theta}_1$ lies somewhere between points D and E in the figure. This requires that the value of diversification be large enough, so that $\widetilde{v}$ is large. In this case the optimal signal is partially informative: it reveals the posterior $\hat{\theta}_1$ with some probability q, and $\overline{\theta}_1$ with the remaining probability $1-q$, where q is such that the expected posterior is equal to the prior, p_1. This is the second case in Corollary 1.

Third, suppose the set S only contains less informative signals, so that $\mathcal{P}_S$ only covers some region where $\nu(\theta_1)$ is concave. Graphically, $\underline{\theta}_1$ and $\overline{\theta}_1$ lie somewhere between points C and D. This again requires that the value of diversification be large enough, so that $\widetilde{v}$ is large enough, and also that the technologies be relatively similar, so that the prior is close to the middle. In this case the concave closure of $\nu(\theta_1)$ over $\mathcal{P}_S$ is equal to $\nu(\theta_1)$—there is no value from information disclosure, and the optimal signal is perfectly uninformative, inducing a posterior equal to the prior. Here, the principal's signals are relatively uninformative, the value of diversification is large enough, and the technologies are symmetric enough, so revealing information would lead to more extreme posteriors, which in equilibrium agents would over-react to, compared to the first-best.

Figure 2 illustrates the 3 cases of Corollary 1, plotting the principal's value function and the posteriors induced by the optimal information design in each case.

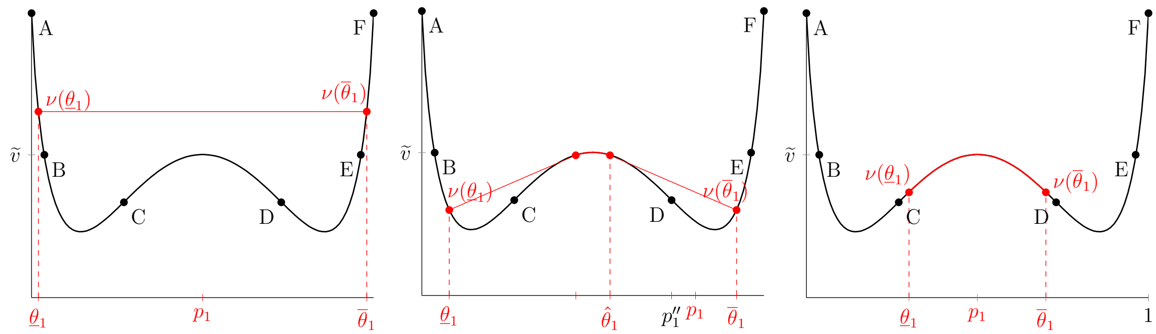

Figure 2: The value function $\nu(\theta_1)$ (in red), with $f(x) = -(1 - x^a)^{\frac{1}{a}}$, with $a = 0.35$.
Case 1: $\underline{\theta}_1 = 0.02, \overline{\theta}_1 = 0.98, p_1 = 0.5$ (left)
Case 2: $\underline{\theta}_1 = 0.08, \overline{\theta}_1 = 0.92, p_1 = 0.8$ (middle)
Case 3: $\underline{\theta}_1 = 0.3, \overline{\theta}_1 = 0.7, p_1 = 0.5$ (right)

Going back to the more general setting, with $N > 2$, Proposition 3 characterizes the optimal signal structure. We can illustrate the principal's objective for $N = 3$, as we did for the $N = 2$ case.

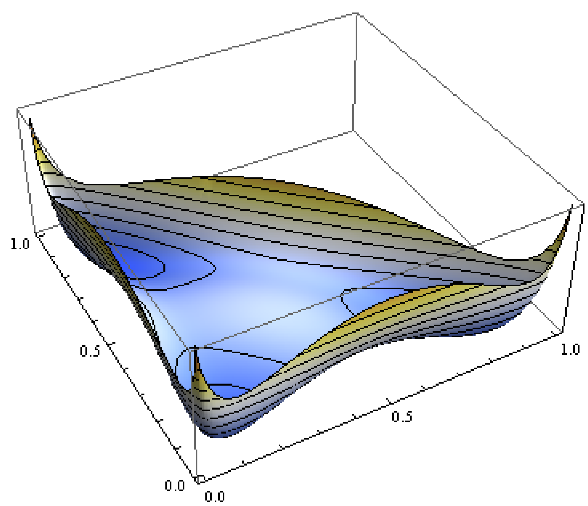

Figure 3: The value function $\nu(\Theta)$, for $N = 3$ and $f(x) = -(1 - x^a)^{\frac{1}{a}}$, with $a = 0.3$.

Figure 3 plots the principal's value function $\nu(\theta_1, \theta_2, \theta_3)$, over the simplex with $N = 3$, where $\theta_3 = 1 - \theta_1 - \theta_2$. Each cross section of this graph, holding one $\theta_i = 0$, is similar to the case with $N = 2$, in Figure 1. In this example, with $f(x) = -(1 - x^{0.3})^{\frac{1}{0.3}}$, the value function is convex towards the vertices, has a concave region around the center of the simplex, and dips between the center and each of the 3 vertices.

5 Conclusion

This paper offers a novel model of variety in contests with technological uncertainty, where agents can pursue different technologies to compete. We study a tractable setting with many agents and an arbitrary number of possible technologies, where we can easily characterize the agents' equilibrium strategies.

We use this framework to study the problem of information disclosure regarding the value of different possible technologies. We fully characterize the optimal signal structure that maximizes the principal's expected payoff from the contest, as a function of the set of all signals available to the designer. We show that the informativeness of the optimal signal structure crucially depends on three main features of the environment: (i) the value of technological diversity; (ii) the quality of the principal's information; and (iii) the extent of technological uncertainty. Each of these factors affects the principal's choice of information structure, as it affects the key trade-off between diversification and focus.

Revealing more precise information about the technologies induces more extreme posteriors, which incentivizes agents to focus on more promising technologies in equilibrium. However, the equilibrium allocation of agents' efforts may over-react to such asymmetries in their beliefs regarding the different technologies, compared to the principal's first-best allocation. Because the technologies are uncertain, the principal's payoff includes the option value of developing less promising technologies, so diversification is also valuable, and conflicts with the incentive to focus on more promising technologies. The optimal signal structure balances these considerations, and can be maximally informative, partially informative, or completely uninformative in different cases.

These results apply to any contest setting where agents can pursue different approaches, such as in procurement, contests for innovation, promotions within organizations, and others. All of these settings have in common the feature that the agents and the principal may be unsure about which technology, idea, or project will be most valuable or feasible ex post.

6 References

Alonso, Ricardo and Odilon Câmara (2016) "Persuading voters," *American Economic Review*, Vol. 106, pp. 3590–3605.

Aoyagi, Masaki (2010) "Information feedback in a dynamic tournament," *Games and Economic Behavior*, Vol. 70, pp. 242–260.

Benkert, Jean-Michel and Igor Letina (2016) "Designing dynamic research contests," *University of Zurich, Department of Economics, Working Paper*.

Bhattacharya, Sudipto and Dilip Mookherjee (1986) "Portfolio choice in research and development," *The RAND Journal of Economics*, pp. 594–605.

Bimpikis, Kostas, Shayan Ehsani, and Mohamed Mostagir (2014) "Designing dynamic contests," *Working paper, Stanford University*.

Bockstedt, Jesse, Cheryl Druehl, and Anant Mishra (2016) "Heterogeneous Submission Behavior and its Implications for Success in Innovation Contests with Public Submissions," *Production and Operations Management*.

Boudreau, Kevin J, Nicola Lacetera, and Karim R Lakhani (2011) "Incentives and problem uncertainty in innovation contests: An empirical analysis," *Management Science*, Vol. 57, pp. 843–863.

Cabral, Luis (1994) "Bias in market R&D portfolios," *International Journal of Industrial Organization*, Vol. 12, pp. 533–547.

Cabral, Luis MB (2003) "R&D competition when firms choose variance," *Journal of Economics & Management Strategy*, Vol. 12, pp. 139–150.

Che, Yeon-Koo and Ian Gale (2003) "Optimal design of research contests," *The American Economic Review*, Vol. 93, pp. 646–671.

Dasgupta, Partha and Eric Maskin (1987) "The simple economics of research portfolios," *The Economic Journal*, Vol. 97, pp. 581–595.

Ederer, Florian (2010) "Feedback and motivation in dynamic tournaments," *Journal of Economics & Management Strategy*, Vol. 19, pp. 733–769.

Fu, Qiang, Jingfeng Lu, and Jun Zhang (2016) "On Disclosure Policy in Tullock Contests with Asymmetric Entry," *Canadian Journal of Economics*, Vol. 49.

Fullerton, Richard L and R Preston McAfee (1999) "Auctioning entry into tournaments," *Journal of Political Economy*, Vol. 107, pp. 573–605.

Gross, Daniel P (2015) "Creativity Under Fire: The Effects of Competition on Creative Production," *Available at SSRN 2520123*.

——— (2017) "Performance feedback in competitive product development," *The RAND Journal of Economics*, Vol. 48, pp. 438–466.

Halac, Marina, Navin Kartik, and Qingmin Liu (2017) "Contests for experimentation," *Journal of Political Economy*, Vol. 125.

Huang, Yan, Param Vir Singh, and Kannan Srinivasan (2014) "Crowdsourcing new product ideas under consumer learning," *Management science*, Vol. 60, pp. 2138–2159.

Kamenica, Emir and Matthew Gentzkow (2011) "Bayesian persuasion," *American Economic Review*, Vol. 101, pp. 2590–2615.

Kireyev, Pavel (2016) "Markets for Ideas: Prize Structure, Entry Limits, and the Design of Ideation Contests," *HBS, Working Paper*.

Klein, Arnd Heinrich and Armin Schmutzler (2016) "Optimal effort incentives in dynamic tournaments," *Games and Economic Behavior*.

Kovenock, Dan, Florian Morath, and Johannes Münster (2015) "Information sharing in contests," *Journal of Economics & Management Strategy*, Vol. 24, pp. 570–596.

Krishnan, Vish and Shantanu Bhattacharya (2002) "Technology selection and commitment in new product development: The role of uncertainty and design flexibility," *Management Science*, Vol. 48, pp. 313–327.

Laclau, Marie and Ludovic Renou (2016) "Public persuasion."

Lemus, Jorge and Guillermo Marshall (2017) "Dynamic Tournament Design: An Application to Prediction Contests."

Letina, Igor (2016) "The road not taken: competition and the R&D portfolio," *The RAND Journal of Economics*, Vol. 47, pp. 433–460.

Letina, Igor and Armin Schmutzler (2017) "Inducing Variety: A Theory of Innovation Contests."

Mathevet, Laurent, Jacopo Perego, and Ina Taneva (2017) "On information design in games,"Technical report, mimeo.

Moldovanu, Benny and Aner Sela (2001) "The optimal allocation of prizes in contests," *American Economic Review*, pp. 542–558.

Taylor, Curtis R (1995) "Digging for golden carrots: an analysis of research tournaments," *The American Economic Review*, pp. 872–890.

Terwiesch, Christian and Yi Xu (2008) "Innovation contests, open innovation, and multiagent problem solving," *Management science*, Vol. 54, pp. 1529–1543.

Toh, Puay Khoon and Taekyu Kim (2013) "Why put all your eggs in one basket? A competition-based view of how technological uncertainty affects a firm's technological specialization," *Organization Science*, Vol. 24, pp. 1214–1236.

Xin, Feng and Jingfeng Lu (2016) "The Optimal Disclosure Policy in Contests with Stochastic Entry: A Bayesian Persuasion Perspective," *Economics Letters*, Vol. 147.

Zhang, Jun and Junjie Zhou (2016) "Information disclosure in contests: A Bayesian persuasion approach," *The Economic Journal*, Vol. 126, pp. 2197–2217.