UTS Business School –
Department of Economics
Dr Chau Chak Wing Building
14 – 28 Ultimo Road, Ultimo NSW 2007

T: +61 2 9514 7713
economics@uts.edu.au

Department of Economics
Working Paper Series

'Diversification and Information in Contests'

Jorge Lemus [1]
Emil Temnyalov [2]

[1] University of Illinois Urbana-Champaign, USA
[2] University of Technology Sydney, Australia

No. 12 / 2019 **July 2019**

Diversification and Information in Contests*

Jorge Lemus† and Emil Temnyalov‡

July 15, 2019

Abstract

We study information disclosure and diversification in contests with technological uncertainty—where agents can pursue different technologies to compete in the contest, but there is uncertainty regarding which is the right one. The principal can credibly reveal information about the technologies to affect the agents' choices. Information revelation may prompt agents to work on the right technology, which is valuable for the principal, but it can also reduce technological diversification, which may be detrimental for the principal in a setting with technological uncertainty. We characterize the optimal information disclosure policy and show that it can be maximally or partially revealing, or completely uninformative, depending on: (i) the value of diversification; (ii) the quality of the principal's information; and (iii) the extent of technological uncertainty. Our results apply to various managerial settings such as innovation contests, tournaments within organizations, and procurement.

JEL codes: O32, C72, D62, D72, D83.

Keywords: information disclosure, contests, variety, diversification.

*For helpful comments we thank seminar and conference participants at APIOC 2018, IIOC 2019, the University of Melbourne, and the University of Toronto, Rotman School of Business.

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

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

1 Introduction

Many contests feature technological uncertainty: there are multiple technologies, approaches, or ideas that contestants could pursue to compete in the contest, and it is not clear which of these will ultimately be successful. This is the case in innovation contests, such as prediction contests hosted on digital platforms (e.g., Kaggle or Innocentive) or contests organized by a firm (e.g., the 2009 Netflix contest soliciting recommendation algorithms). Another example are agricultural yield contests, such as those sponsored by the National Corn Growers Association, where farmers compete to produce the highest yield. Farmers can use different technologies, but their yield is subject to uncertainty (e.g., the performance of a technology may depend on the weather or the soil). Technological uncertainty is also present in procurement contests, such as those hosted by the U.S. Department of Defense, where contestants may not fully know the preferences of the procurer over different possible attributes or designs that they could develop. Similarly, technological uncertainty is also present in contest-like settings within organizations: workers face uncertainty about the impact of different tasks on their chances of a promotion.

Uncertainty and competition affect how the agents choose technologies in the contest. Some agents will work on technologies that turn out to be unsuccessful, so developing multiple approaches increases the principal's chances to ultimately implement a successful technology. In such a setting, the principal may be able to reveal some information regarding the different technologies in order to affect the agents' beliefs about the different alternatives, which in turn changes the equilibrium allocation of effort across technologies in the contest. For instance, in prediction contests, the contest sponsor provides data that contestants use to train and develop their algorithms; these data reveal information about the probability of success of different algorithms, and hence the contestants can use it to choose which approach to develop. In yield contests, the principal can reveal the performance of agricultural approaches used in the past, although this is not a guarantee that they will work well the future, because weather conditions could be different. Similarly, in procurement contests participants submit designs and prototypes that can be evaluated by the procurer at some preliminary stage; if the procurer reveals these evaluations to the contestants, then it will affect the contestants' beliefs about the likelihood of success of each design, and hence also the contestants' choices of what to develop subsequently.

Revealing too much information, however, can be detrimental for the principal, because it may bias the agents. In prediction contests, for example, the principal wants to avoid "over-fitting," i.e., algorithms that perform well in one dataset, but not in general. Thus, the principal seeks to incentivize the contestants to develop general algorithms which will be valuable beyond the existing data. In procurement, the procurer may want to not reveal some of its evaluations, in order to avoid steering the contestants into developing very similar designs. More broadly, in any contest with technological uncertainty, revealing some information about the different possible technologies affects the contestants' choices and it may induce too little diversity of technologies. On the other hand, in order to avoid competition, in some settings the agents might prefer a more diversified portfolio of technologies, relative to the principal. Hence information disclosure regarding the probability of success of different technologies, approaches, or ideas, can be a valuable contest design tool for the principal in the presence of technological uncertainty and competition.

Our model explores technological uncertainty in contests—where uncertainty creates an option value for diversification across technologies—and we study the principal's optimal information disclosure policy. This optimal policy balances the trade-off between providing the agents with better information about the technologies and the option value of diversification. In our model, agents compete in a contest and choose one among multiple possible technologies, one of which is the correct choice—e.g., it will be implemented ex post by the principal. Neither the agents nor the principal are informed ex ante about which technology is the correct one.[1] The principal designs a public experiment, and commits to its result, in order to reveal information to the agents regarding the value or likelihood-of-success of each technology.[2] Information disclosure improves the agents' beliefs about the technologies, but it may induce an equilibrium where too many agents choose the more promising technology and there is too little technology diversification from the perspective of the principal.

Our results apply to a number of managerial problems in innovation, procurement, and organizations, where technological uncertainty is important. The model sheds light on the information design aspect of online platforms such as Kaggle, DrivenData,

[1]These different technologies may represent different approaches to solve a problem in the case of innovation, different characteristics, features, or designs in the case of procurement, or different tasks or projects in the case of a worker competing within an organization.

[2]This is, we study optimal information disclosure in a Bayesian persuasion framework.

and crowdAI, among others, which offer firms the possibility to sponsor contests to outsource their data science needs. In these contests, when contestants choose among multiple technologies to develop their algorithms (e.g. machine learning, regression methods, or other prediction algorithms) they do not know which method will turn out to be the best. However, contestants can infer the performance of different alternatives from testing them on a dataset provided by the principal, who chooses how much data to disclose.[3] The contest sponsored by Netflix to improve its recommendation system is another example where the information disclosure affected the agents' choices. 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 the potential of these different approaches. 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 important contest design question is how much of their available data to reveal in the contest. 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.[4]

One more instance where information disclosure affected diversification across technologies is the Hyperloop Pod Competition, sponsored by Space X. The goal of this competition was to test vehicle-prototypes for the Hyperloop. The competition consisted of several stages and the results were made publicly available at the end of each stage. 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. Was it optimal for the contest designer to disclose this information after the first round? One possible drawback is that some teams may be tempted to abandon their approaches and switch to the most promising approach after

[3] In these contests, the contest designer partitions the available data into a "test dataset" and an "evaluation dataset." The test dataset allows participants to learn how their algorithms perform, but the final prize allocation is based on the performance of the algorithm over the evaluation dataset.

[4] Indeed, during the Netflix contest one user followed an approach that performed well at a preliminary stage, and this information was revealed to all participants, which led this approach to be widely adopted by other competitors, per https://www.thrillist.com/entertainment/nation/the-netflix-prize.

the first stage, 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 the best ex post—the contest designer may see value in a variety of approaches.

We characterize the optimal information structure that maximizes the principal's expected payoff from the contest. The main trade-off is one between diversification and focus: revealing more precise information induces agents to focus on more promising technologies, but if the agents 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 the 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 is the right one. Hence developing multiple different approaches has an option value: even if one technology looks 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, which is related to the measure of risk-aversion associated with the principal's objective function, 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, i.e., an experiment that reveals with very high probability which technology is the right one, then she is more likely to want to reveal such information. In practice, however, the principal may not have access to 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.

Related Literature. Our paper contributes to the recent literature on diversification in contests. Letina and Schmutzler (2017) characterize the optimal prize structure when the designer wants to induce a variety of approaches. Our paper offers a new model of variety, and 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. 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. The existing 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, whereas Mihm and Schlapp (2018) study the optimal information disclosure to maximize the provision of effort when players are uncertain about the principal's preferences. Kovenock et al. (2015) study 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, 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) show 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) explores 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, in the framework of 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), Ederer et al. (2018), 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. This 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.

2 Model

A continuum of agents, indexed by $i \in [0, 1]$, compete in a contest that awards a single prize V to the agent that develops the "best" solution to a problem.[5] There are N different "technological approaches" to solve the problem, and only one of these approaches is the correct one, i.e., is the approach that will deliver a feasible solution that will be implemented ex post. Each agent has one indivisible unit of effort to allocate to one of the technologies. This assumption allows us to focus on the allocation of effort

[5] In prediction contests (such as those hosted on Kaggle) there is a well-defined notion of the best solution. A typical measure of performance includes the sum of square errors, but the right algorithm, the one that minimizes this metric in the evaluation dataset, is ex ante unknown. Using the test dataset to evaluate different algorithms provides a noisy measure of an algorithm's final performance.

across technologies, rather than on the intensity of effort exerted by the agents. Agents hold a common-prior $p = (p_j)_{j=1}^N$, where $p_j \in [0, 1]$ is the belief that technology j is the correct approach, and $\sum_{j=1}^{N} p_j = 1$.

We model competition among agents in a reduced form. Given that a mass of x_j agents allocate their effort towards technology j, the probability that one of these agents wins the contest is uniform. In other words, we assume that the prize allocation in the contest is anonymous: if two agents both choose the ex post successful technology, they have the same chances of winning the contest. Agents that choose an ex post unsuccessful technology never win the contest. Hence, the probability of an individual winning the contest is uniform over the agents who choose the right technology.

Assumption 1. *Let $j \in N$ be the correct ex post approach. If a measure of x agents work on technology j, each of them has a probability (density) $s(x) = \frac{1}{x}$ of winning the contest. Agents that work on technology $k \neq j$ never win the contest.*

Agents' Payoffs. Under Assumption 1, when $j \in N$ is the correct approach ex post, and the distribution of agents over technologies is $x = (x_1, ..., x_N)$, agent i's payoff from choosing technology k is

$$u_i(k) = \begin{cases} \frac{V}{x_j} & \text{, if } k = j \\ 0 & \text{otherwise} \end{cases}.$$

For any belief $p = (p_j)_{j=1}^N$, agent i's expected payoff from choosing technology k, given the distribution $x = (x_1, ..., x_N)$ of all the other agents over the technologies is

$$E[u_i(k)|p] = \frac{p_k V}{x_k}.$$

Principal's Payoff. The principal's payoff depends on how many agents allocate their effort to the ex post successful technology. If the distribution of agents over technologies is $x = (x_1, ..., x_N)$ 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. This function represents the gains from agents' efforts (or investments) into the chosen technology. Having more agents develop the correct technology improves the principal's payoff, but there are decreasing marginal returns to agents' efforts.[6] An ex post unsuccessful technology (i.e., not feasible or not valuable) results in unproductive or wasteful effort, from the designer's perspective. Thus, for a given belief $q = (q_1, ..., q_N)$ of the principal over the likelihood of success of each technology, the principal's expected payoff from allocation x is

$$E[v(x)|q] = \sum_{j=1}^{N} q_j f(x_j) - V.$$

Since V is not influenced by information design (i.e., it is separable with respect to the beliefs), in the reminder of the analysis we suppress this term in the principal's payoff.

For any interior beliefs $q = (q_j)_{j=1}^N$, the first-best allocation of agents solves the following problem:

$$x^{FB} \in \arg\max_{x \in [0,1]^N} \sum_{j=1}^{N} q_j f(x_j) \quad \text{subject to} \quad \sum_{j=1}^{N} x_j = 1. \tag{1}$$

Proposition 1 (First best allocation)**.** *When $f(\cdot)$ is increasing, differentiable and concave, for any interior beliefs $q = (q_j)_{j=1}^N$ about the technologies, the solution to Problem (1) is characterized by*

$$\frac{f'(x_i^{FB})}{f'(x_j^{FB})} = \frac{q_j}{q_i} \quad \forall i, j \in N.$$

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, because the first best allocates agents so as to equalize the marginal expected gains from the technologies. Given that f is concave (f' is decreasing), the principal allocates more agents to technologies that are more promising and diversifies: agents are allocated across multiple technologies, because it is ex ante uncertain which technology is ex post implementable. Note that the value of diversification for the principal is measured by the concavity of f.

Next, we consider the equilibrium of the game where agents choose which technology to work on. Here, they take into account both competition with other agents (a crowding-out effect) and also the likelihood that a technology is successful. More promising technologies attract more agents, but competition pushes agents to also work on less promising technologies.

[6] This can be seen as a reduced-form representation of a model where there is a finite number of agents that draw scores out of a distribution, and the principal's payoff is the maximum score.

Definition 1. *A distribution of agents over technologies* $x = (x_1, ..., x_n)$ *is an equilibrium if* $\sum_{j=1}^{N} x_j = 1$ *and no agent* i *who allocated effort towards technology* k *can improve its payoff by allocating its effort towards technology* $\ell \neq k$*.*

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

Proposition 2 (Equilibrium characterization)**.** *With a common interior belief* Θ *about the technologies, the equilibrium mass of agents working on technology* j *is*

$$x_j = \theta_j.$$

Proof. Consider an allocation of agents where a mass x_j of 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 expected payoff which implies

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

This condition is equivalent to $x_j = \frac{\theta_j}{\theta_k} x_k$, 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 $x_k = \theta_k$, for all k. □

Proposition 2 shows that in equilibrium more agents are allocated to more promising technologies, 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. For this reason, the equilibrium allocation is generally inefficient and features too much or too little diversification, relative to the first-best allocation.

2.1 Equilibrium inefficiency: over- and under-reaction.

Consider the case of two technologies. Let $\Theta = (\theta, 1 - \theta)$ be the common belief about the two technologies, i.e., agents believe that technology 1 is the correct choice with probability θ, and assume without loss of generality that $\theta \geq \frac{1}{2}$. Let x^E be the equilibrium mass of agents allocated to technology 1, and let x^{FB} the the first-best allocation of agents to technology 1. From Proposition 1 and Proposition 2 we have $\theta f'(x^{FB}) = (1 - \theta) f'(1 - x^{FB})$ and $x^E = \theta$.

What is the relation between x^E and x^{FB}? When $\theta = \frac{1}{2}$ the equilibrium is efficient: $x^E = x^{FB} = \frac{1}{2}$. But when $\theta \neq \frac{1}{2}$, in equilibrium agents may "under-react" or "over-react" to information about technology 1; that is, in equilibrium too few or too many agents may work on the more promising technology, relative to the first best. Whether agents over- or under-react will depend on the shape of the principal's payoff at the current belief in relation to the competitive force induced by the contest.

To characterize more generally when agents over- or under-react to information about the ex-ante most promising technology in equilibrium, we define

$$H(\theta) = \theta f'(\theta) - (1-\theta) f'(1-\theta).$$

Note that this function is 2-fold symmetric around $\theta = \frac{1}{2}$, and $H(0) \leq 0$ and $H\left(\frac{1}{2}\right) = 0$.

Proposition 3. *For $\theta > \frac{1}{2}$ (w.l.o.g.), agents over-react to information about technology 1, i.e. $x^E > x^{FB}$, if $H(\theta) < 0$, and they under-react, i.e. $x^E < x^{FB}$, if $H(\theta) > 0$.*

Whether agents under or over react to information is related to the coefficient of relative risk aversion associated with f:

$$r_f(\theta) \equiv \frac{-\theta f''(\theta)}{f'(\theta)}.$$

Proposition 4. *When the principal is not too risk averse, agents under-react to information about the most promising technology in equilibrium. Specifically, when*

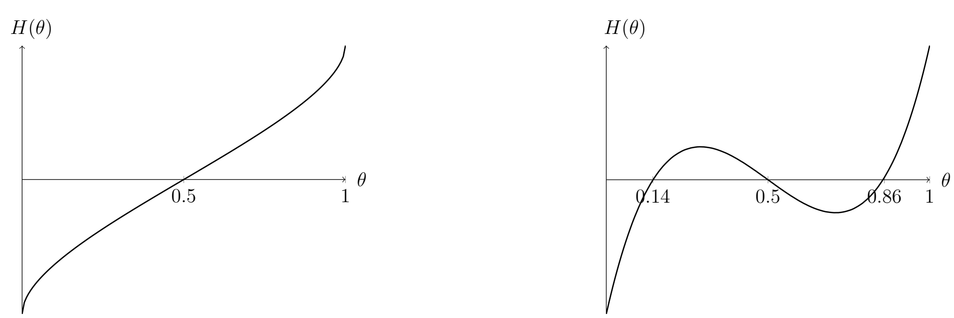
$$(1-\theta) r_f(\theta) + \theta r_f(1-\theta) < 1, \qquad \textit{for all } \theta \in [0,1],$$

too few agents allocate to the most promising technology in equilibrium, relative to the principal's optimal allocation.

Consider for example $f(x) = x^a$ for some $a \in (0,1)$. In this case, $r_f(x) = 1 - a < 1$ so the principal is not too risk averse, and Proposition 4 implies that agents under-react to information about the most promising technology. Intuitively, the principal would like more agents working on the most promising technology, but competition between agents pushes them to work on a more diverse set of technologies. This is easy to see in the limit case when $a \to 1$, so $f(x) = x$, and the principal would allocate *all* agents to the most promising technology. However, this is not an equilibrium because an agent would rather work on any of the other technologies and face no competition. Thus, relative to the first best allocation, in equilibrium there is too much diversification.

When the condition in Proposition 4 is violated, i.e., when the principal's risk aversion is relatively high, we can still determine whether agents in equilibrium react or under react to information. To illustrate this point, consider $f(x) = 1 - \exp(-\lambda x)$. We have $r_f(x) = \lambda x$ and thus $(1-\theta)r(\theta) + \theta r(1-\theta) = 2\lambda\theta(1-\theta)$. The parameter λ corresponds to the (absolute) risk aversion of the principal. When $\lambda < 2$ we can apply Proposition 4 to show that agents will under-react to the most promising technology. However, when $\lambda > 2$, the principal is too risk averse, and agents over-react to the most promising technology for $\theta \in \left[\frac{1}{2}, \frac{1}{2} + \theta^*\right]$ and under-react for $\theta > \frac{1}{2} + \theta^*$. Intuitively, when the belief is $\theta \in \left[\frac{1}{2}, \frac{1}{2} + \theta^*\right]$, the high risk aversion of the principal implies that the principal would prefer diversification, but in equilibrium too many agents work on the most promising technology. Therefore, agents may over or under-react to information.

Figure 1 illustrates these two cases. The left panel shows the function $H(\theta)$ when the principal's payoff is defined by $f(x) = x^a$, with $a = 0.5$. In this case, in equilibrium, too few agents allocate to the most promising technology: there is always under-reaction to information about the most promising technology. The right panel shows $H(\theta)$ when the principal's payoff is defined by $f(x) = 1 - \exp(-\lambda x)$, with $\lambda = 2.5$. In this case, in equilibrium, too few agents allocate to the most promising technology when $\theta \in [0, 0.14) \cup (0.86, 1]$, and too many agents allocate to the most promising technology when $\theta \in (0.14, 0.86)$.



(a) $H(\theta)$ when $f(x) = \sqrt{x}$. **(b)** $H(\theta)$ with $f(x) = 1 - \exp(-2.5x)$.

Figure 1: Function $H(\theta)$ for different preferences represented by $f(\cdot)$.

Given that generically competing agents' incentives and the principal's preferences are misaligned, we explore whether the principal can design an experiment to persuade agents to work towards different technologies. Intuitively, the principal can disclose information in such a way as to exploit the over- or under-reaction of agents in equilibrium to achieve a better allocation of agents.

3 Information design

We now turn to the analysis of optimal information disclosure. In general, the principal may be able to reveal information to the agents regarding the feasibility or likelihood of success of different technologies, for example by revealing the results of preliminary testing or evaluations of different alternatives, or some data that the agents can use to form beliefs about the alternatives. Should the principal reveal such information? In a situation where technology-specific prizes are not contractible, can the contest designer use information revelation strategically to improve the equilibrium allocation of agents across technologies? We study this question in a Bayesian persuasion framework: we assume the principal can ex ante commit to an information disclosure policy, i.e., the principal designs a public experiment that will reveal public information about the technologies to all agents. For instance, in prediction contests this experiment corresponds to the size of the test and evaluation datasets. In innovation contests, this experiment corresponds to revealing the result of prototype testing.

An experiment is a signal structure $s = (M, \widetilde{G}(\cdot|j))$, where M is a set of messages and $\widetilde{G}(m|j)$ is the probability that message $m \in M$ is sent when the state of nature j, i.e., when technology j is the correct choice ex post. Let S be the set of all such signal structures available to the principal. Importantly, we do not assume that the principal has access to *every* signal structure. This is motivated by the observation that in many practical applications a perfectly informative signal is impossible to generate: in many applications it is unfeasible to design an experiment that eliminates all the uncertainty. Additionally, the problem is trivial if a perfectly informative signal exists: the principal would immediately reveal to the agents which technology is the correct one and all agents would work on this technology. Instead, we solve for the optimal information disclosure policy for *any arbitrary* set of available signals. The set S allows for discrete or continuous posteriors, and it allows for both partitional or noisy signal structures.

Each signal structure $s \in S$ induces some distribution over posterior beliefs, $G_s(\Theta)$, and we denote the set of posterior beliefs 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 posterior beliefs that can be induced by some signal.

Lemma 1. *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.*

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. In particular, when there are only 2 technologies, the set of feasible posteriors is an interval.

The value of information disclosure can be analyzed in terms of the posterior beliefs that a signal structure induces for the agents and the principal. Recall that when agents hold the belief Θ, in equilibrium we have $x_j = \theta_j$. Therefore, the principal's expected payoff from inducing a posterior Θ is

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

As in Kamenica and Gentzkow (2011), the value of information disclosure is described by the convexity of $\nu(\Theta)$. Let $\hat{\nu}(\Theta, \mathcal{P}_S)$ be the concave closure of $\nu(\Theta)$ over $\mathcal{P}_S$. The principal strictly benefits from persuasion whenever $\hat{\nu}(\Theta_0, \mathcal{P}_S) > \nu(\Theta_0)$ around the prior Θ_0. We can easily characterize whether $\nu(\cdot)$ is concave or convex.

Lemma 2. *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 $j = 1, ..., N$. □

Lemma 2 implies that we need to study the concavity of $\theta f(\theta)$ to determine whether or not there are gains from persuasion. This can be characterized very intuitively as follows:

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

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. This implies that when $r_f(\theta)$ is large, i.e. the principal is very risk averse, the principal does not benefit from information disclosure. And when $r_f(\theta)$ is low, i.e. the principal is not too risk averse, the principal benefits from information disclosure.

In subsection 2.1 we showed that the relative coefficient of risk aversion also captures the value of diversity to the principal relative to the equilibrium level of diversity. Hence, the trade-off between diversity and information disclosure is captured by r_f. We now illustrate how the value of information disclosure relates to this coefficient for the case of two technologies. Consider again $f(x) = x^a$, with $a \in (0,1)$. In this case, $r_f(x) = 1 - a < 2$ for all $x \in [0,1]$, and from Lemma 2 we get that the function $\nu(\cdot)$ is globally convex, so there are always gains from information disclosure, for any prior and for any $\mathcal{P}_S$. Recall that for this function agents always under-react to good news, i.e., $x^E < x^{FB}$ for $\theta > \frac{1}{2}$, where x^E is the equilibrium mass of agents working in technology 1, and x^{FB} is the efficient mass of agents working on that technology. Hence the two observations—that agents under-react to news and that it is optimal for the principal to reveal as much information as possible—are indeed closely connected.

On the other hand, for $f(x) = 1 - \exp(-\lambda x)$, the function $\nu(\cdot)$ can be concave around the middle and convex near the extremes—for λ 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. Whether the principal can benefit or not from revealing information thus depends on the prior and on $\mathcal{P}_S$: information disclosure could benefit the principal if the prior is relatively extreme and the principal can induce very informative signals; otherwise, disclosure is not optimal in the concave region of $\nu(\Theta)$, where agents would over-react to news. Figure 2 shows the value for the principal of inducing posterior θ for $f(x) = x^{0.5}$ (shown in the left panel) and for $f(x) = 1 - \exp(-6x)$ (shown in the right panel).

(a) $\nu(\theta)$ when $f(x) = x^{0.5}$. **(b)** $\nu(\theta)$ when $f(x) = 1 - \exp(-6x)$.

Figure 2: Function $\nu(\theta)$, the value for the principal of inducing posterior θ, for different preferences for the principal, which are represented by $f(\cdot)$.

Intuitively, whether there are gains from information disclosure or not (i.e. whether $\nu(\cdot)$ is locally convex or concave) depends on the value the principal assigns to diversification across technologies. When $r_f(x)$ is relatively large, i.e. $f(\cdot)$ is very concave, diversification is more valuable to the principal than to the agents, because the equilibrium allocation, 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. In contrast, when $r_f(x)$ 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 illustrated in the example with $f(x) = x^{0.5}$ in Figure 2a.

Under mild conditions, the shape of $\nu(\cdot)$ with $N > 2$ technologies is analogous to the two cases shown in Figure 2. The next lemma characterizes the shape of $\nu(\cdot)$.

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

(i) all of its global maxima are at the vertices of the $N - 1$ simplex;

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

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 $\nu(0, ..., 0, 1, 0, ..., 0) = 1 \cdot f(1)$. The only points that can obtain the global maximum of $f(1)$ are the vertices of the simplex. To see this, consider any point $\Theta' = (\theta'_1, ..., \theta'_N)$ such that $\nu(\Theta') = \sum_j \theta'_j f(\theta'_j) = 1$. Then $f(\theta'_j) \geq f(1)$ for some $j \in N$. Since f is increasing, this requires $\theta'_j = 1$, hence the point Θ' is a 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.

The center of the simplex is always a critical point of $\nu(\Theta) = \sum_{j=1}^{N} \theta_j f(\theta_j)$ subject to $\sum_{j=1}^{N} \theta_j = 1$ because the FOC with respect to θ_j, $f(\theta_j) + \theta_j f'(\theta_j) = \lambda$, for $j = 1, ..., N$, and $\sum_{j=1}^{N} \theta_j = 1$ is clearly satisfied at $\theta_j = \frac{1}{N}$ for all $j \in N$. If this is the unique solution, then that critical point is a local minimum, and $\nu(\Theta)$ must be convex everywhere, since all of its global maxima are at the vertices, by part (i). 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. 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 the 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. $\square$

Lemma 3 implies some useful properties of the principal's objective function. Notice that $\text{sign}[2f'(x) + xf''(x)] = \text{sign}[2 - r_f(x)]$, so we can interpret this condition in terms of the relative risk coefficient: if $r_f(x) < 2$, the function $\nu(\cdot)$ is convex; if $r_f(x)$ crosses 2 only once, then there is at most one interior local maximum.[7]

[7] A sufficient condition for $r_f(\cdot)$ to be monotone would be to assume that the sign of f''' is constant: when $f'''(x) > 0$ ($f''' < 0$), r_f is decreasing (increasing).

Additionally, Lemma 3 shows that $\nu(\cdot)$ has global maxima with value $f(1)$ at the extreme points of the simplex, because at those extreme points the principal and agents know with certainty which technology is the correct choice (it will be implemented ex post). Thus, every agent in equilibrium works on that technology, which is the optimal allocation for the principal. This immediately implies that if the principal has access to a perfectly informative signal, revealing that signal is 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.

Also, Lemma 3 shows that 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.[8] Whether such a concave region exists or not depends on the functional form of the production function, $f(x)$. Lemma 3 provides a sufficient single-crossing condition to ensure that the objective has at most one interior local maximum: if $2f'(x) + xf''(x)$ crosses 0 at most once in $[0, \frac{1}{N}]$, then the objective either only has an interior local minimum, or it has a unique interior local maximum at the center.

We assume that f satisfies the properties characterized by Lemma 3(ii) for the remainder of the analysis, noting that the characterization of an optimal signal structure does not critically rely on this assumption—one can 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.

Optimal information design. We now characterize the optimal signal structure in our setting. Let Θ_0 be the common prior belief over the N technologies. Conditional on that prior belief there exists a set of posteriors $\mathcal{P}_S(\Theta_0)$ that the principal can induce that are consistent with Bayes' rule. In particular, we always have that $\Theta_0 \in \mathcal{P}_S(\Theta_0)$, because the principal can always decide not to reveal anything, so the posterior equals the prior.

Let $\widetilde{v} \equiv \sup\{\nu(\Theta) : \nu''(\Theta) \leq 0, \quad \Theta \in \Delta^N\}$ be the largest value of the value function over the region where it is concave in the simplex $\sum_{j=1}^{N} \theta_j = 1$.[9] Let $\Theta_C \equiv \{\Theta :$

[8] In this case there would also be N local minima between the vertices and the center.

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

$\hat{\nu}(\Theta, \mathcal{P}_S) = \nu(\Theta)\}$ be the set of all posteriors where the value function ν agrees with its concave closure over $\mathcal{P}_S$, i.e., where $\nu = \hat{\nu}$. Denote by $\partial\mathcal{P}_S$ the boundary of $\mathcal{P}_S$. Finally, recall that $\mathcal{P}_S$ is a convex subset of the simplex Δ^N. We can now characterize the optimal disclosure policy.

Proposition 5. *The optimal disclosure policy s^* is*

1. ***maximally informative*** *if $\nu(\overline{\theta}) \geq \tilde{v}$ for all $\overline{\theta} \in \partial\mathcal{P}_S$; then s^* induces a distribution over posterior beliefs with support consisting only of points in the boundary of the feasible set of posteriors, with distribution G_s s.t. $\mathrm{E}_{G_s}[\Theta] = \Theta_0$.*
2. ***partially informative*** *if $\nu(\overline{\theta}) \geq \tilde{v}$ and $\mathcal{P}_S \nsubseteq \Theta_C$; then s^* induces a distribution with support consisting of boundary points in $\partial\mathcal{P}_S$ and in $\mathcal{P}_S \cap \Theta_C$, with distribution G_s s.t. $\mathrm{E}_{G_s}[\Theta] = \Theta_0$.*
3. ***uninformative*** *if $\mathcal{P}_S \subseteq \Theta_C$; then s^* induces the prior, $\Theta = \Theta_0$.*

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 for a given prior; and (iii) the extent of technological uncertainty. We illustrate the intuition for each of these features in the case of two technologies. First, if ν is globally convex, then the optimal signal structure is always maximally informative and the optimal persuasion experiment reveals results that lead to extreme posterior beliefs. On the other hand, in the case where the principal is risk averse enough, i.e., when ν is not convex (see Figure 2), the optimal information design problem is more subtle, so we focus on this case in the example below.

The special case with $N = 2$. Consider an arbitrary set of posteriors $\mathcal{P}_S$ that the principal can induce (note that this must be a convex set). There exists $\underline{\theta}_1$ and $\overline{\theta}_1$ with $0 \leq \underline{\theta}_1 \leq \overline{\theta}_1 \leq 1$, such that any posterior belief over technologies $\Theta = (\theta_1, 1 - \theta_1) \in \mathcal{P}_S$ is characterized by $\theta_1 \in [\underline{\theta}_1, \overline{\theta}_1]$; hence we identify $\mathcal{P}_S = [\underline{\theta}_1, \overline{\theta}_1]$.

Figure 3 shows the function $\nu(\theta) \equiv \nu(\theta, 1 - \theta)$ for the case $f(x) = 1 - \exp(-\lambda x)$, as in Figure 2b, highlighting three features: (1) the inflexion points of $\nu(\cdot)$, C and D, at $\theta = \frac{2}{\lambda}$ and $\theta = 1 - \frac{2}{\lambda}$, respectively; (2) the unique interior maximum at $\theta = \frac{1}{2}$; and (3) the points B and E where the value of $\nu(\cdot)$ equals its value at the interior local maximum. The function ν is concave in the region $\left[\frac{2}{\lambda}, 1 - \frac{2}{\lambda}\right]$ and convex otherwise.

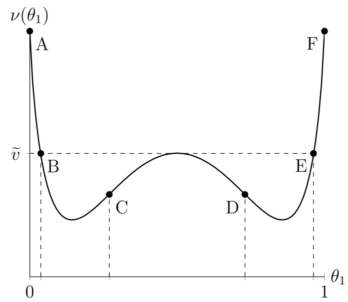


Figure 3: The value function $\nu(\theta_1)$ with $f(x) = 1 - \exp(-\lambda x)$, with $\lambda = 6$.

We have that $\widetilde{v} = \sup\{\nu(\theta) : \nu''(\theta) < 0\} = f\left(\frac{1}{2}\right) = 1 - \exp(-\lambda/2)$. We can restate Proposition 5 for this special case:

Corollary 1. *Suppose, without loss of generality, that the prior is $\Theta_0 = (\theta_0, 1-\theta_0)$ with $\theta_0 \geq 0.5$. Also for the sake of exposition, suppose that $\underline{\theta}_1 = 1 - \overline{\theta}_1$, so the set of posterior beliefs the principal can induce is symmetric with respect to 0.5.*

The optimal disclosure policy s^* is

1. ***maximally informative** if $\nu(\overline{\theta}_1) \geq \widetilde{v}$ (see Figure 4 left panel); s^* induces a binary posterior distribution $(\underline{\theta}_1, q; \overline{\theta}_1, 1-q)$ with $q\underline{\theta}_1 + (1-q)\overline{\theta}_1 = \theta_0$.*
2. ***partially informative** if $\nu(\overline{\theta}_1) < \widetilde{v}$ and $\theta_0 \in (\frac{2}{\lambda}, \overline{\theta}_1)$ (see Figure 4 middle panel); s^* induces a binary posterior distribution $(\hat{\theta}_1, q; \overline{\theta}_1, 1-q)$ with $q\hat{\theta}_1 + (1-q)\overline{\theta}_1 = \theta_0$, where $\hat{\theta}_1 = \sup\{\theta : \hat{\nu}(\theta, [\underline{\theta}_1, \overline{\theta}_1]) = \nu(\theta)\}$.*
3. ***uninformative** if $\nu(\overline{\theta}_1) < \widetilde{v}$ and $\overline{\theta}_1 \leq \frac{2}{\lambda}$ (see Figure 4 right panel); s^* induces a degenerate posterior distribution, $(\theta_0, 1)$.*

Figure 4 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.

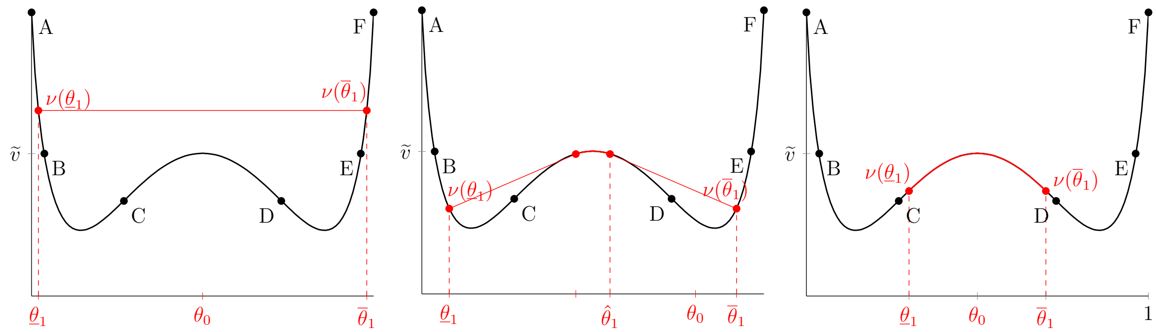


Figure 4: The value function $\nu(\theta_1)$ (in red), with $f(x) = -(1-x^a)^{\frac{1}{a}}$, with $a = 0.35$.
Case 1: Feasible posteriors $\mathcal{P}_S = [\underline{\theta}_1, \overline{\theta}_1]$ with $\underline{\theta}_1 = 0.02$ and $\overline{\theta}_1 = 0.98$. Prior $\theta_0 = 0.5$ (left)
Case 2: Feasible posteriors $\mathcal{P}_S = [\underline{\theta}_1, \overline{\theta}_1]$ with $\underline{\theta}_1 = 0.08$ and $\overline{\theta}_1 = 0.92$. Prior $\theta_0 = 0.8$ (middle)
Case 3: Feasible posteriors $\mathcal{P}_S = [\underline{\theta}_1, \overline{\theta}_1]$ with $\underline{\theta}_1 = 0.3$ and $\overline{\theta}_1 = 0.7$. Prior $\theta_0 = 0.5$ (right)

The principal's optimal signal structure depends on the a priori asymmetry of the technologies (i.e. on the prior Θ_0), the quality or informativeness of the feasible signals the principal can use for a given prior (i.e., $\mathcal{P}_S(\Theta_0)$), and on the value of diversification to the principal (i.e. on $\widetilde{v}$).

First, if S includes signals that are highly informative, so that the principal can induce posterior beliefs in $\mathcal{P}_S$ including posterior beliefs close enough to 0 and 1, then maximal disclosure is optimal, and 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. In this case, the concave closure of $\nu(\cdot)$ is the 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, θ_0. This is illustrated in the left panel of Corollary 1.

Similarly, when the value of diversification to the principal is relatively low, then $\widetilde{v}$ is low and we also have $\max\{\nu(\underline{\theta}_1), \nu(\overline{\theta}_1)\} > \widetilde{v}$, in which case the optimal signal is also maximally informative, as in the first case of Corollary 1. The reason is that when the value of diversification to the principal is low (i.e. the $r_f(x)$ coefficient is low enough), revealing information to the agents increases the principal's expected value because in equilibrium agents "under-react" to asymmetries in the technologies, when beliefs are extreme. Thus, the principal benefits from inducing extreme posteriors. This leads to an optimal disclosure rule that mixes between the 2 most extreme posteriors possible within $\mathcal{P}_S$. Graphically, when $f(\cdot)$ is relatively less concave and the coefficient $r_f(\cdot)$ is lower, then the value of $\widetilde{v}$ decreases, 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 so that $\nu(\underline{\theta}_1) < \widetilde{v}$ and $\nu(\overline{\theta}_1) < \widetilde{v}$. Moreover, suppose that the technologies are ex-ante asymmetric, with technology 1 being more likely ex ante, with θ_0 to the right of point D. Graphically, in the middle panel of Corollary 1, $\underline{\theta}_1$ 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. This is the second case in Corollary 1.

Third, suppose the set S only contains less informative signals, so that the principal can induce beliefs in $\mathcal{P}_S$ only around 0.5 in the region where $\nu(\cdot)$ is concave. Graphically, in the right panel of Corollary 1, $\underline{\theta}_1$ and $\overline{\theta}_1$ lie somewhere between points C and D. 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. The principal is constrained by the set of signals it can use to persuade the agents, which are relatively uninformative signals. 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.

4 Conclusion

When there are different approaches to tackle a problem, and agents compete in a contest to find the correct solution, we ask whether it is beneficial for the contest sponsor (the principal) to disclose information regarding the different approaches. We find that it is not always beneficial to reveal that one technology is more promising than the rest when the principal cares about diversification: revealing information can induce too many agents to work on the most promising technology, which reduces diversification.

We present a tractable framework to study contests with technological uncertainty and to analyze the trade off between information revelation and diversification. In our setting, each agent chooses one out of N available technologies to compete in the contest, and only one of these technologies is ex post the correct one. The principal can commit to reveal to the agents the results of an experiment that signals the success of each technology. 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.

5 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.

Ederer, Florian, Richard Holden, and Margaret Meyer (2018) "Gaming and strategic opacity in incentive provision," *The RAND Journal of Economics*, Vol. 49, pp. 819–854.

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.

Mihm, Jürgen and Jochen Schlapp (2018) "Sourcing innovation: On feedback in contests," *Management Science*, Vol. 65, pp. 559–576.

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.

A Appendix: Proofs

Proof of Proposition 3

Proof. Suppose that technology 1 is the ex-ante most promising technology, so the prior is $\theta > \frac{1}{2}$. When $\theta f'(\theta) > (1-\theta)f'(1-\theta)$ we have that $x^E = \theta < x^{FB}$, because f' is decreasing (by the concavity of f). Thus, in this case, agents in equilibrium *under-react* to information about the ex-ante most promising technology: too few agents are working on technology 1 relative to the optimal allocation for the principal. Analogously, when $\theta f'(\theta) < (1-\theta)f'(1-\theta)$, we have that $x^{FB} < x^E = \theta$, so agents in equilibrium *over-react* to information about the ex-ante most promising technology, i.e., too many agents work on technology 1 relative to the optimal allocation for the principal. Over- or under-reaction would be determined analogously in the case when $\theta < \frac{1}{2}$, where $\theta f'(\theta) > (1-\theta)f'(1-\theta)$ implies over-reaction, while $\theta f'(\theta) < (1-\theta)f'(1-\theta)$ implies under-reaction. □

Proof of Proposition 4

Proof. It is immediate to see that $H\left(\frac{1}{2}\right) = 0$. When $f(0) = 0$ and $f' > 0$, by concavity we have

$$f(0) \leq f(\theta) + f'(\theta)(0-\theta) \Rightarrow \theta f'(\theta) \leq f(\theta)$$

This implies that $\lim_{\theta\to 0} \theta f'(\theta) \leq \lim_{\theta\to 0} f(\theta) = 0 < f'(1)$, so $H(0) \leq 0$.

Consider the following auxiliary result:

Lemma 4. *Suppose that H is differentiable. Then for any θ s.t. $H(\theta) = 0$,*

$$sign\ H'(\theta) = sign\ [1 - (1-\theta)r_f(\theta) - \theta r_f(1-\theta)].$$

Proof. Taking derivative we have:

$$\begin{aligned} H'(\theta) &= f'(\theta) + \theta f''(\theta) + f'(1-\theta) + (1-\theta)f''(1-\theta) \\ &= f'(\theta)\left[1 + \frac{\theta f''(\theta)}{f'(\theta)}\right] + f'(1-\theta)\left[1 + \frac{(1-\theta)f''(1-\theta)}{f'(1-\theta)}\right] \end{aligned}$$

Consider a point $\theta \in (0,1)$ such that $H(\theta)=0$. Then, $f'(\theta)=\frac{(1-\theta)}{\theta}f'(1-\theta)$ so

$$\begin{aligned} H'(\theta) &= \frac{f'(1-\theta)}{\theta}\left\{(1-\theta)\left[1+\frac{\theta f''(\theta)}{f'(\theta)}\right]+\theta\left[1+\frac{(1-\theta)f''(1-\theta)}{f'(1-\theta)}\right]\right\}. \\ &= \frac{f'(1-\theta)}{\theta}\left[1-(1-\theta)r(\theta)-\theta r(1-\theta)\right]. \end{aligned}$$

Given that f is strictly increasing ($f'>0$) and that $\theta>0$ we have the result. □

Therefore, because $H(0)\leq 0$ and $H(\frac{1}{2})=0$, we can analyze the sign of H under some regularity conditions.

Corollary 2. *Suppose that $(1-\theta)r_f(\theta)+\theta r_f(1-\theta)<1$ for all $\theta\in[0,1]$. Then, $H(\cdot)$ is strictly increasing, implying that in equilibrium equilibrium agents under-react to information about technology 1, i.e., $x^E<x^{FB}$ for $\theta>\frac{1}{2}$.*

Proof. If $1>(1-\theta)r(\theta)+\theta r(1-\theta)$ for all θ, the function H cannot cross zero before $\theta=1/2$, because if it crosses zero, it must cross from below. Since $H(0)<0$, this is impossible. In this case, H only crosses zero at $\theta=1/2$ and from below. Thus, $H(\theta)<0$ for $\theta<1/2$ and $H(\theta)>0$ for $\theta>1/2$, which means that agents under-react to good news. □

□

Proof of Lemma 1

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. □

Proof of Proposition 5

Proof. The proofs follows from the concavification argument.

1. A maximally informative signal obtains when the set of posteriors is rich enough, so the then the global maxima of $\nu(\Theta)$ over $\mathcal{P}_S$ is below the concave closure of $\nu(\Theta)$ over the region $\mathcal{P}_S$, because $\delta\mathcal{P}_S$ includes points towards the vertices of the simplex, so the concave closure of $\nu(\Theta)$ over $\mathcal{P}_S$ corresponds to the plane that connects the boundary of $\mathcal{P}_S$.

Then the optimal signal s^* only induces posteriors in $\partial\mathcal{P}_S$, so for any arbitrary prior $p \in \mathcal{P}_S$, Bayesian consistency of the posteriors determines the distribution over posteriors on $\partial\mathcal{P}_S$.

2. A partially informative signal obtains when the set of posteriors is limited, so the concave closure of $\nu(\Theta)$ over $\mathcal{P}_S$ coincides with $\nu(\Theta)$ for some values (near the center of the simplex Δ^N).

3. An uninformative signal obtains when the set of posteriors is concentrated towards the center of the simplex Δ^N, where ν is concave, so the concavification of $\nu(\cdot)$ over $\mathcal{P}_S$ and $\nu(\cdot)$ itself coincide. □