Disincentive Effects of Evaluation*

Raphael Boleslavsky†
Bruce I. Carlin‡
Christopher S. Cotton§

February 26, 2019

ABSTRACT

In a model of project design, evaluation, and selection, we explore how the incentives to improve the design of projects depend on the availability of funding and the process of evaluation. We show that project designers (researchers or NGOs) prefer to subject their projects to less-rigorous evaluations than donors or funding agencies would prefer, ex post. We also find that increases in both funding availability and the informativeness of evaluations tend to decrease investments in project quality. By implication, increased availability of funding or more-informative evaluations can lead to the implementation of fewer high-value projects.

*This paper has benefited from insightful discussions with several international development practitioners who have provided helpful insights into the evaluation and funding processes within their organizations. Cotton is grateful for financial support provided through his position as the Jarislowsky-Deutsch Chair in Economic and Financial Policy at Queen's University.

†University of Miami, Department of Economics, r.boleslavsky@miami.edu.

‡University of California, Los Angeles, Anderson School of Management, bruce.carlin@anderson.ucla.edu.

§Corresponding Author. Queen's University, Department of Economics, cotton@econ.queensu.ca.

1 Introduction

Pilot studies and early-stage evaluations provide insights into the likely effectiveness and impact of research projects and social programs ahead of full scale or long-term implementation. Insights gained through data collection and analysis can help improve the design and implementation of these projects. Such evaluations also provide institutions, donors, or investors with better information when they award funding or select projects to implement more broadly.

Recent advances in data collection and evaluation technologies increase the feasibility of rigorous evaluations during a program's pilot stage or initial period of the program.[1] At the same time, growing expectations from donors and the general public for financial accountability have led many institutions to increase their reliance on evidence when choosing programs to fund, implement, or expand.[2]

The information generated by such evaluations is generally believed to lead to the design and implementation of higher-quality projects. From this perspective, the more informative the evaluation, the better. At the same time, however, conducting a careful, rigorous evaluation of a project is costly and may be limited by institutional or logistical constraints. It may not be feasible to evaluate certain projects, and the evaluation of others may face resistance from the NGOs, project managers, or researchers who may be hurt by evidence showing that their efforts have been ineffective. From the conventional perspective, how rigorous of an evaluation to conduct is generally seen as a trade-off between the likely benefits of being better informed about a project and the costs, both material and agency, of conducting the evaluation.

Our analysis adds some nuance to the conventional perspective. We present a styled model of project design, evaluation, and selection that builds on recent insights from work on strategic evidence production in competitive environments. The model provides several insights regarding the impact of funding availability and evaluation rigor on the incentives to invest in higher quality proposals, the availability of high value opportunities, and the effectiveness of funding.

First, we show that having the capacity to fund more projects can *reduce* the number of

[1] It is easier to conduct randomized evaluations at a local or regional level, for example, with platforms such as Tangerine and Kobo Toolbox making it straightforward for field enumerators to complete learning evaluations and surveys in the field using their mobile devices. It is also becoming less costly to collect, store and share administrative data on program performance, and to analyze data.

[2] Major philanthropic foundations such as the Bill & Melinda Gates Foundation and the MasterCard Foundation emphasize research and evaluation as being central to their missions. The Gates Foundation (2019), for example, states "Our aim is to integrate evaluation into the fabric of our work." The MasterCard Foundation (2015) emphasizes the role of evaluation in faciliatating "evidence-based decisions and amplified impact." Additionally, there has been a recent shift in social and international development sectors to focus more on "results-based financing" of projects through instruments such as social or development impact bonds and payment-for-results agreements between governments or donors and program managers. By their very nature, these mechanisms tie together evaluation and funding outcomes. See, for example, Boddild-Jones and Gustafsson-Write (2019) and Gustafsson-Wright, Gardiner and Putcha (2015).

high value projects that are implemented. Limited funding incentivizes competition among those seeking funding for their projects, leading them to invest more in developing higher quality proposals, which are more likely to develop into high value projects. When funding is widely available, in contrast, there is less investment in project quality and fewer high-value projects may be implemented in equilibrium.

Second, we show that those responsible for developing and designing the projects and programs under review often prefer less-rigorous evaluations than those responsible for the allocation of funding. A more-rigorous evaluation more likely to discover that a project is of low value. Consequently, with a more rigorous evaluation the project developers are more likely to lose out on funding opportunities, an undesirable outcome which they would like to avoid.

Third, we explore the interaction between evaluations and agent incentives to invest in designing better projects. we show that if project developers anticipate more-rigorous evaluations, they may invest *less* in their projects and put forward lower-quality opportunities. In turn, this negatively affects the overall quality of the projects vying for funding.

To explore these issues, we compare three different environments. In the first, there are no evaluations. In the second, the agents who seek funding (e.g. NGOs, institutions or researchers) are responsible for designing the process evaluations of their own project proposals. In the third, there are fully-revealing evaluations of all proposals. Although requiring full evaluation of all proposals ensures that funding is efficiently allocated based on the projects' realized quality, it also leads to less investment by the agents when developing projects in the first place. We find that the disincentive effects generally dominate the efficiency effects. Consequently, requiring fully informative evaluations leads to the development and funding of fewer high-value projects.

To put our results into a specific context, consider the case of alternative non-governmental organizations (NGOs) looking for funding to expand their program models which are designed to increase school-completion rates among at-risk girls in a developing country. A western development agency (e.g. DFID, USAID) or the local-country government pledges to fund the broad implementation of one of the NGO's program, and will choose the program that is expected to have the greatest impact. In competition for the funding opportunity, the NGOs can work to improve their program models. This may involve taking additional steps to ensure that the program design addresses the perceived needs within the local context, that it incorporates the most-promising techniques, that a plan is in place for reaching the individuals and communities that will benefit most, and that the NGO engages a team with the necessary qualifications, experience, and local connections to ensure smooth implementation. With design and implementation plans for the program models in place, pilot studies may be undertaken within a subset of communities to establish the viability and likely impact of the alternative programs. Our results show that the NGOs prefer to subject their proposed program models to less-rigorous evaluation than the funding agency or government does. This is true even when conducting an evaluation is costless for the NGOs. But, this does not mean that it is obvious that funding organizations should require higher standards of evaluation. This is because our results also show that requiring more-rigorous evaluations will lead to less initial investment by the NGOs when they design their programs. This can ultimately lead to the proposal, selection, and implementation of less-effective programs and decrease the impact of the funding on school completion rates.

Similar insights apply in the allocation of research funding. The National Institute for Health (NIH), for example, often issues funding over multiple rounds, first funding a pilot study, which is evaluated before funding for the full research proposal is allocated. Our results suggest that although the pilot studies are useful for improving the efficient allocation of funding across alternative proposals, such pilots may have the unintended consequence of reducing the incentive of researchers to invest in developing careful, promising research designs and proposals.

It is widely known that impact evaluations are costly, sometimes face institutional resistance, and suffer from quality control issues. At the same time, it is also generally assumed that if we can overcome these barriers and conduct high-quality, rigorous evaluations of projects and proposals, then governments and donors will be able to better compare opportunities, which will in turn lead to the funding of more high-value opportunities. We show how better evaluations may have precisely the opposite effect, driving down the investment in developing quality proposals and ultimately leading to the funding of fewer high-value opportunities. The analysis is intended as a first step in developing a more-complete understanding of the benefits and costs of higher-quality evaluations.

2 Literature

A substantial literature discusses the benefits of pilot studies and impact evaluations, and the merits of evidence-based policy more generally. These papers generally argue that better evaluation leads to the design and selection of more-effective projects, programs and policies, but also that evaluations face barriers in terms of feasibility, costs, and institutional resistance. Gertler, et al. (2016) provides a detailed review of the benefits and best practices of impact evaluation. Duflo and Banajee (2011) and Karlan and Appel (2011) review the evaluation of several past international development projects, arguing that such evaluations help implementers design better policies and more efficiently allocate funding. In other fields, Crosswaite and Curtice (1994) argue that pilot studies increase accountability to better justify the use of funds and van Teijlingen and Hundley (2001) argue that pilots offer many benefits including identifying failures ahead of project implementation. Head (2010) argues that the use of information and evidence in policymaking is often limited by misaligned preferences and entrenched commitments make organizations resistant to change. Mebrahtu (2002) and Merchant-Vega (2011) argue that organizations may be resistant to evaluations that show their work as being ineffective. This stream of past work on the costs and benefits of evaluation focuses on case studies, qualitative assessments and intuition. We complement this literature by developing a game theoretic model of project design, evaluation, and selection in order to better understand the relationships between funding availability, evaluation, and project design, which we use to rigorously evaluate some of this literature's arguments.

The game theoretic model that we develop builds on recent work that considers how later-stage evaluations affect incentives in earlier stages of strategic environments. For example, it is well-established in the career concerns literature that more monitoring can discourage effort (e.g. Dewatripont, et al 1999 and Holmstrom 1999). Similarly, Coate and Loury (1993) and Taylor and Yildirim (2011) consider how a decision maker's access to information about an agents type affects the agents decision to invest in quality. In the contracting literature, it is also well established that a principal may want to commit to imperfect monitoring technologies, as doing so leads to more favorable actions taken by agents (e.g. Cremer 1995, Sappington 1986, 1991). Along this same line, our work contributes to a large body of work that implies that people may be better off ignoring information or by committing not to collect it (e.g. Hirshleifer 1971; Morris and Shin 2002).[3]

More closely related to our work are several papers that build on this literature to consider settings in which agents are not only concerned about assessments on an absolute scale, but are also concerned about their assessments relative to other agents. Boleslavsky and Cotton (2015a) show that politicians competing in an election prefer to run on less-moderate platforms when they anticipate that more information about their quality will emerge during a campaign. Boleslavsky, Cotton and Gurnani (2017) show that firms competing in a market set less competitive prices as consumers become better able to distinguish between products. Boleslavsky and Cotton (2015b) show that schools competing to place students invest less to provide high-quality education when employers are better able to evaluate the quality of their individual graduates. The underlying model in our current paper is most similar to the model in Boleslavsky and Cotton (2015b) in that it explicitly models both investment in quality and the design of an evaluation technology (e.g. grading policy vs. impact evaluation). However, the earlier work does so in a two-agent environment which is not suited to study the allocation of a large amount of funding across a large number of projects.[4] Furthermore, by extending the earlier work to consider a continuum of heterogeneous agents, we are able to consider how the availability of funding affects project design and evaluation strategies, issues that could not be addresed in a model with two agents.

3 Model

There is a continuum of agents with mass 1. Each individual agent is responsible for developing and proposing an "opportunity" (e.g. project, program, policy) to a decision maker, who selects a subset of opportunities to receive funding.

[3]See Bikhchandani, Hirschleifer, and Welch (1992), Teoh (1997), Angeletos and Pavan (2007), and Amador and Weill (2012), among others.

[4]A number of other papers endogenize the design of evaluation technology in competitive environments without considering investment decisions. See, for example, Gentzkow and Kamenica (2017), Au (2018), Au and Kawaii (2018), Boleslavsky and Cotton (2018), Li and Norman (2018), Hulko and Whitmeyer (2018).

First, each agent simultaneously chooses how much to invest in order to improve the initial design of his or her own opportunity, which affects the expected effectiveness of the opportunity if it receives funding. After the agents develop proposals, each proposal is evaluated. The evaluation process produces a more precise estimate of each proposal's likely effectiveness. Following the evaluation process, the decision maker updates her beliefs about the likely effectiveness of each proposal, accounting for the initial quality of the proposal and any evidence revealed during the evaluation. The decision maker then awards funding to the opportunities with the highest expected effectiveness.

The following timeline illustrates the order of the game:

$t = 0$	$\rightarrow$	$t = 1$	$\rightarrow$	$t = 2$
Agents develop proposals choosing quality q_i		Proposals evaluated with informativeness γ_i		Decision maker allocates funding

At $t = 0$, each agent $i \in [0, 1]$ chooses the quality of his proposal, which we denote as $q_i \in [0, 1]$. Proposal quality q_i determines the *probability* that agent i's opportunity will deliver a high value if it is implemented. The decision maker and the agents are capable of observing the quality of each proposal. Therefore, all players anticipate that proposal i is high value ($\tau_i = H$) with probability q_i and low value ($\tau_i = L$) with probability $1 - q_i$. Investment level q_i is associated with a convex cost for agent i: $C(q_i) = q_i^2/2\rho_i$. Differences in parameter $\rho_i > 0$ captures variation in agent ability, experience, and the merit of the initial idea. The higher ρ_i, the lower i's marginal cost of improving proposal quality. The values of ρ_i in the population are uniformly distributed on $[0, 1]$.

After the agents simultaneously and independently choose their proposal qualities, all proposals undergo evaluations, which reveal additional information about the value that will be generated by the proposals if they are funded. Formally, the evaluation process is modeled as a random variable, R_i, whose distribution depends on the proposal type τ_i. A realization, r_i, of this random variable represents the outcome of the evaluation. We assume that R_i takes on two values, either H or L, where

	$\Pr(R_i = H)$	$\Pr(R_i = L)$
High value ($\tau_i = H$)	1	0
Low value ($\tau_i = L$)	$1 - \gamma_i$	γ_i

The informativeness of the evaluation R_i is represented by γ_i. Such an evaluation clearly reveals a low-value project with probability γ_i. That is, when the realization of the evaluation is low, $R_i = L$, the decision maker learns for sure that the opportunity is low value. When the realization is $R_i = H$, however, some uncertainty remains about project value (except in the case of $\gamma_i = 1$). The more informative the evaluation—the higher is γ_i—the greater the likelihood that the evaluation identifies a low value project. Consequently, the decision maker is more confident that a project with a high evaluation, $R_i = H$ is truly high value. The posterior belief that a project with a high evaluation is actually high quality is

$$g_i(q_i, \gamma_i) = \frac{q_i}{q_i + (1 - \gamma_i)(1 - q_i)}. \tag{1}$$

When $\gamma_i = 0$, the evaluation reveals no additional evidence about proposal i and the decision maker's beliefs about project value are determined only by the quality of the proposal, and $g_i = q_i$. When $\gamma_i = 1$, the evaluation is fully revealing about the value of opportunity i, and $g_i = 1$.

Following the evaluation and updating of beliefs about proposal effectiveness, the decision maker awards funding to a measure $s \in (0, 1)$ of the proposed projects. An agent who's proposal is selected by the decision-maker receives a benefit normalized to one, giving total utility $u_i = 1 - q_i^2/(2\rho_i)$. Agents who do not receive the benefit of having their projects implemented receive total utility $u_i = -q_i^2/(2\rho_i)$. The decision-maker receives a payoff normalized to 1 for each high-value project and 0 for each low-value project that receives funding, with an average payoff equal to the mass of high-value projects that receive funding.[5] Given this, the decision-maker's expected benefit of implementing proposal i is simply the posterior belief that $\tau_i = H$, given the available information (equal to 0 when $r_i = L$, and equal to g_i when $r_i = H$). The decision maker strictly prefers to select the measure s projects with the highest posterior probability of being high value.

Within this framework, we consider three alternative assumptions involving the design of evaluations:

(i) **No evaluations**, which is equivalent to the case of fully-uninformative evaluations, where $\gamma_i = 0$.

(ii) **Fully-revealing evaluations**, which involve $\gamma_i = 1$ and are preferred by the decision-maker at $t = 2$.

(iii) **Agent-designed evaluations**, where we assume that agents strategically design γ_i at $t = 2$ as they compete for funding.

In the second and third cases, we assume that the choice of evaluation rigor, $\gamma_i \in [0, 1]$ has no cost. By abstracting from the costs of evidence production, we can focus on the setting where the decision to produce less-informative evidence is driven by strategic considerations rather than cost considerations. This assumption clarifies our contribution to the literature on evaluations, by isolating a novel strategic force that limits informativeness, abstracting from the issue of costly evaluations that has already received considerable attention.

We focus on Perfect Bayesian Equilibria of the game.

4 Analysis

4.1 Equilibrium with No Evaluations

When there are no evaluations, the decision-maker allocates funding to the measure s agents with the highest quality proposals. In equilibrium, there exists a proposal-quality threshold

[5] A low-value project is one with benefits that do not exceed the costs of implementation. A high-value project offers benefits that exceed implementation costs.

at which relatively high-ability agents (as parameterized by ρ_i) develop proposals. The threshold is sufficiently high as to be out-of-reach of lower ability agents for whom the costs of proposal development are higher. Furthermore, if the equilibrium threshold is less than 1, then it must not to be in the interest of any agent to invest a bit more than the threshold. In this case, the probability of funding for projects at the threshold must be 1. If the threshold equals 1, then no higher investment is possible. In this case, the probability of funding at the threshold may be less than 1. Moreover, the measure of funded projects at the threshold must be s. In this way, if an agent invests a bit less than the threshold, he will guarantee that his project will not receive funding.[6]

There are two cases to consider, depending on the availability of funding, s, relative the average ability of agents to develop proposals ($E[\rho_i] = 1/2$). We describe these cases in Proposition 1.

Proposition 1. *(Equilibrium with No Evaluations) When $\theta_i = 0$ for each i, equilibrium investment depends on the availability of funding.*

i. When $s \leq 1/2$, agents invest

$$q_i = \begin{cases} 1 & if \quad \rho_i \geq 1/(1+2s) \\ 0 & if \quad \rho_i < 1/(1+2s). \end{cases} \tag{2}$$

ii. When $s > 1/2$, agents invest

$$q_i = \begin{cases} \sqrt{2(1-s)} & if \quad \rho_i \geq 1-s \\ 0 & if \quad \rho_i < 1-s. \end{cases} \tag{3}$$

When funding is relatively scarce compared to the average ability to develop proposals, competition for the limited funding generates fierce competition. In this case, the measure of agents who develop the *maximum* possible quality exceeds the measure of available funding, s. In this case, some rationing occurs at the maximum level of quality (assuming $s < 1/2$).

When more funding is available—relative to agents' expected ability—competition between agents is less fierce and agents develop lower-quality proposals. In this case, the s highest-ability agents invest in proposals that are just high enough quality that any lower ability agents find it too costly to develop such proposals. The s highest-ability agents all develop proposals of equal quality and each receives funding.

The decision maker's payoff is the measure of high value projects implemented. Thus, the decision maker is concerned with maximizing the social benefits generated by the funding level, s.[7] When funding is relatively scarce, proposals are assured to be high value, and

[6] Remember that agents are infinitesimal, and a deviation by a single agent does not change the measure of funded projects at the threshold.

[7] In the analysis, only high value projects provide a net benefit to society. The assumption that low-value projects return a benefit of 0 to the decision maker effectively assumes that the benefits of such projects for society equal their costs of implementation.

therefore the decision maker is assured of implementing s high-value projects in equilibrium, with $u_d = s$. When funding is relatively abundant, the decision maker payoffs are $u_d = s\sqrt{2(1-s)}$.

If proposal quality were exogenous, the decision maker would weakly prefer to have more funding.[8] However, in our environment, proposal quality is chosen by the agents, who tend to produce lower-quality proposals as the available funding increases. In this case, the decision maker does not strictly prefer higher funding (even when that funding comes at zero cost).

Corollary 1. *(More funding is not strictly better) When there are no evaluations, the number of high value projects that receive funding (which equals u_d) is maximized at $s = 2/3$.*

4.2 Equilibrium with Fully-Revealing Evaluations

Next, we consider the case of fully-revealing evaluations, where $\gamma_i = 1$ for each i. Such an evaluation strategy is always preferred by the decision-maker at time $t = 2$, taking as given the quality of each proposal, q_i. It is commonly viewed as the "gold standard" for interim-stage evaluation, perfectly revealing the value of each project and assuring that the decision maker has the information he or she needs to allocate funding optimally.

Proposition 2 describes the equilibrium of the game when proposals are subject to fully-revealing evaluations. As in the preceding case, the incentives that agents have to invest in their proposals depend on the capacity for funding, s, relative to the average ability of the agent population to develop high value projects, $E[\rho] = 1/2$.

Proposition 2. *(Fully-Revealing Evaluations)*

i. When $s \leq 1/2$, agents each invest $q_i = \rho_i\sqrt{2s}$.

ii. When $s > 1/2$, agents each invest $q_i = \left(1 - \sqrt{1 - 2(1-s)}\right)\rho_i$.

In contrast to the preceding case, where funding was allocated based on each agent's investment level, with fully revealing evaluations the decision maker no longer awards funding based on q_i directly. Rather, she awards funding based on the information about actual project value that is revealed by the evaluation. In this case, even lower ability agents have an incentive to invest a positive amount to develop proposals as part of equilibrium, hoping that an evaluation shows that their project actually has high value.

In this case, agent investments are maximized when $s = E[\rho_i] = 1/2$. This level of funding maximizes competition between agents, given their average ability level.[9] When funding is more scarce than this level, a discouragement effect arises: agents exert less effort because there are fewer prizes, relative to the ability of their competitors. The lower that s is, the greater the discouragement effect and the lower the effort of each agent. When

[8] This is because the net benefits of implementing a high-value projects is positive and the net benefit of implementing low-value projects is assumed to be zero.

[9] With a more-general distribution of ρ, one can show that the corresponding funding level is $s = E[\rho]$.

funding is more plentiful, agents anticipate that the availability of funding will be sufficient to support even some low-value proposals. This also reduces agents' incentives to invest in their projects.

The decision maker's payoff is determined by the measure of high-value projects funded. When funding is relatively scarce ($s \leq 1/2$), the decision maker's payoff is determined by his capacity for funding, $u_d = s$. When funding is relatively abundant ($s > 1/2$), her payoff depends on the number of developed projects that are high-value, with

$$u_d = \frac{1}{2}\left(1 - \sqrt{1 - 2\,(1-s)}\right). \tag{4}$$

Decision maker payoffs are maximized at the same point that agent investment in quality is maximized, i.e. at $s = 1/2$.

Again, we find that being able to fund more projects is only better for the decision maker up to the point. Beyond this point, the decision maker is strictly worse off from having the ability to fund additional projects. We compare the payoff of the decision-maker across the three environments in Section 4.4.

4.3 Equilibrium when agents design evaluations

The decision-maker selects the measure s of opportunities that she believes are most likely to be high value. Since there is imperfect information, there will now exist a threshold belief G such that each agent only receives funding if the posterior belief about his proposal's quality surpasses the threshold: $g_i \geq G$.

As in the no-evaluations case, there are two possibilities. If $G < 1$, then the measure of projects with posterior beliefs at G must equal s in equilibrium.[10] If $G = 1$, then projects must be proven to be high value in order to receive funding, and there may not be enough funding for all high-value projects. This case reduces to the case of fully-revealing evaluations studied above.

Proposition 3. *(Equilibrium when Agents Design Evaluations)*

i. When $s \leq 1/2$, each agent invests $q_i = \rho_i\sqrt{2s}$ and designs fully-revealing evaluations.

ii. When $s > 1/2$, each agent invests

$$q_i = \begin{cases} \frac{\rho_i}{\sqrt{2(1-s)}} & if \quad \rho_i < 2(1-s) \\ \sqrt{2(1-s)} & if \quad \rho_i \geq 2(1-s). \end{cases} \tag{5}$$

[10] If $G < 1$, then any any project with posterior beliefs $g_i = G$ must be accepted in equilibrium. Otherwise agents would have an incentive to marginally increase θ_i in order for positive reports to result in a posterior belief regarding their project just above G. Doing so would lead to only a marginal decrease in the probability of a favorable evaluation and a measurable increase in their probability of funding (increasing to probability 1).

Agents choose evaluation informativeness

$$\gamma_i = \begin{cases} \frac{1-\sqrt{2(1-s)}}{1-q_i} & if \quad \rho_i < 2(1-s) \\ 0 & if \quad \rho_i \geq 2(1-s). \end{cases} \tag{6}$$

According to Proposition 4, when resources are sufficiently scarce (i.e., $s < 1/2 = E[\rho]$), agents conduct fully-revealing evaluations as part of the equilibrium. With scarce funding, agents anticipate that only the projects that are certain to be high value will receive funding. They subject their proposals to fully-revealing evaluations recognizing that in equilibrium that is a necessary condition of funding. In this case, outcomes are the same as if the decision maker designs evaluations herself at $t = 2$.

The more-interesting case from our perspective involves $s > 1/2 = E[\rho]$. With more-abundant funding, agents recognize that even projects with weaker evidence in their favor can receive funding. When designing their evaluations, agents face a trade-off between the probability with which their evaluation generates favorable evidence, and the beliefs associated with favorable evidence. By reducing the informativeness of his evaluation (decreasing γ_i), an agent makes it less likely that their evaluation reveals a low-value project and more likely that the evaluation produces a favorable report ($r_i = H$). At the same time, reducing γ_i means that a favorable report is less convincing, and more-likely associated with a low-value project. In the equilibrium with sufficient funding capacity, the agents prefer the least-informative evaluation that they can design while ensuring that their evaluation is not so uninformative that even a favorable report does not receive funding.

Within the case where $s > 1/2$, as funding capacity increases, the lowest-ability agents (as measured by ρ_i) invest more in their projects while higher-ability agents invest less. Furthermore, given any level of investment, an agent chooses a less-informative evaluation design, on average.

The decision maker's payoffs are again determined by the measure of high-value projects that receive funding. When $s \leq 1/2$, agents choose fully revealing evaluations and the decision maker's payoff is the same as in the fully-revealing evaluation case studied above: $u_d = s$. When $s > 1/2$, funding is more abundant, and the decision maker's payoff depends on both the availability of high-value projects and her ability to identify them. Her payoff simplifies to:

$$u_d = s\sqrt{2(1-s)}, \tag{7}$$

In the case of fully-revealing evaluations, we showed that a decision maker may be better off with limited funding capacity. Being endowed with the ability to implement more projects can be detrimental when it leads to less investment in quality by those competing for funding. This is also the case when agents design evaluations, except that the maximum u_d is no longer achieved when $s = \mu = 1/2$, but rather when $s = 2/3$.

4.4 How funded project value depends on evaluation design

In this section, we compare the effectiveness of funding, determined by the number of high value projects funded, under no evaluations, fully-informative evaluations and agent-designed evaluations. The primary purpose of this analysis is to explore whether evaluations, in particular more-informative evaluations, lead to better outcomes.

As we discussed before, the analysis abstracts from the costs of evaluation. Therefore, the only adverse effects of (more-informative) evaluations involves how the informativeness of evaluations affects agent incentives to develop higher-quality projects.

For any $s \in (0,1)$, the equilibrium of the game with no evaluations features fewer agents developing proposals. However, among developed proposals average quality is higher than in either game with evaluations. Because higher quality projects tend to be developed in the absence of evaluations, introducing evaluations does not necessarily lead to the funding of higher-value projects.

First, consider the case with relatively scarce funding ($s \leq 1/2$). With no evaluations, agents invest enough to ensure that their proposals provide high value. Therefore, evaluations are not needed to identify and fund high value opportunities. In contrast, in both environments with evaluations, agents invest less in developing high quality proposals, but evaluations allow the decision maker to distinguish high and low value projects making up for the lower investment quality. In all three environments, the decision maker funds only high-value projects.

Next, consider the case with more abundant funding ($s > 1/2$). With no evaluations, agents invest more than in the other cases, but not enough to assure that their proposals are high value. Thus, without evaluations, the decision maker ends up funding some low value projects. But, the funding of some low value projects also happens with evaluations. In the case with fully-revealing evaluations, this is because investment is not high enough to develop a sufficient number of high value projects. In the case with agent-designed evaluations, this is because agents choose less-informative evaluations which leave some uncertainty about project value. We find that the environments with no evaluations and agent-designed evaluations result in a larger measure of high-value projects being funded than the environment with fully-revealing evaluations.

Proposition 4. *(Evaluation Design and Decision Maker Payoffs)*

i. When $s \leq 1/2$, in all three environments, the decision maker implements only high value projects and $u_d = s$.

ii. When $s > 1/2$, decision maker's payoff is lower with fully revealing evaluations than with either no evaluations or agent designed evaluations. The decision maker's payoff in the no-evaluations environment is equal to the decision maker's payoff in the agent-designed evaluations environment.

The proposition shows that the decision maker tends to allocate funding to fewer high-value projects when evaluations are fully revealing than when there are no evaluations or when agents are allowed to conduct less-informative evaluations of their projects. Although fully-revealing evaluations improve the ability of the decision maker to distinguish high-value and low-value projects, such evaluations also reduce the incentives of agents to invest in developing high-value projects in the first place. The disincentive effect of fully-revealing evaluations dominate the informational benefits when funding is relatively abundant.

5 Conclusion

In this paper, we developed a stylized model of project design, evaluation and selection to explore how agent incentives to invest in improving the design of their projects depends on the availability of funding and the presence of evaluations.

The analysis produces three main results.

First, the ability to fund more projects may decrease the number of high-value projects that are implemented. This is because the greater availability of funding can decrease agent investments to develop high-value projects.

Second, agents prefer to subject their projects to less-informative evaluations than a funding agency or donor (decision maker) prefers ex post. Noisy evaluations are less likely to show when a project is low value, which agents may prefer in equilibrium as they compete for funding.

Third, the existence and informativeness of evaluations does not increase the number of high value projects that are funded. We show that evaluations can discourage investments in developing higher quality projects. Requiring that proposals are evaluated in a manner that reveals their value leads to fewer high value projects being implemented than when there are no evaluations or when agents can design less-informative evaluations.

There are several benefits to evaluation that are not present in our paper, and therefore our results should be interpreted with care. For example, there are many settings in which agents have little control over the expected value of their projects during a proposal process. These settings are better represented by a model in which proposal quality is fixed and evaluation is strictly beneficial for the effective allocation of funding. Additionally, the evaluations in our framework have no impact on project design. In many environments, however, evaluation can help improve the design of projects, policies and programs, improving long-term outcomes. These considerations are absent in our framework.

With the preceding qualifications in mind, we show that there may be unintended consequences in the way that the individuals or organizations designing and implementing programs, policies or projects react when they anticipate that their proposals face greater scrutiny. Evaluations are generally assumed to facilitate the selection and funding of higher-value projects, and we highlight the possibility that they have the opposite impact. We do not intend our results to discourage the use of impact evaluation, or to suggest that improvements to evaluation technology are detrimental. But, our results do suggest that the the impact of evaluations on outcomes is not as straightforward as previously assumed. Our analysis is intended as a first step in providing a more-complete understanding of the trade-offs that exist when weighing of costs and benefits of conducting pilot studies and impact evaluations for individual projects.

References

Amador, Manuel, and Pierre-Olivier Weill. 2012. Learning from Private and Public Observations of Others Actions. *Journal of Economic Theory* 147: 910-940.

Angeletos, George-Marios, and Alessandro Pavan. 2007. Efficient Use of Information and Social Value of Information. *Econometrica* 75: 1103-1142.

Au, Pak Hung. 2018. Competition in Designing Pandora's Boxes. *Mimeo.*

Au, Pak Hung, Keiichi Kawai. 2018. Competitive Information Disclosure by Multiple Senders. *Games and Economic Behavior*, forthcoming.

Banerjee, Abhijit V. and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty.* New York: Public Affairs.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 100: 992-1026.

Boggild-Jones, Izzy and Emily Gustafsson-Wright. 2019. *A global snapshot: Impact bonds in 2018.* Washington, DC: Brookings Education Plus Development.

Boleslavsky, Raphael, and Christopher Cotton. 2015a. Information and Extremism in Elections. *American Economic Journal: Microeconomics* 7: 165-207.

Boleslavsky, Raphael, and Christopher Cotton. 2015b. Grading Standards and Education Quality. *American Economic Journal: Microeconomics* 7: 248-279.

Boleslavsky, Raphael, and Christopher Cotton. 2018. Limited capacity in project selection: competition through evidence production. *Economic Theory* 65: 385-421.

Boleslavsky, Raphael, Christopher Cotton and Haresh Gurnani. 2017. Demonstrations and price competition in new product release. *Management Science* 63: 2016-2026.

Coate, Stephen, and Loury, Glenn. 1993. Will Affirmative-Action Policies Eliminate Negative Stereotypes?. *American Economic Review* 83: 1220-1240.

Cremer, J. 1995 Arms Length Relationships. *Quarterly Journal of Economics* 110: 275-295.

Crosswaite and Curtice. 1994. Disseminating research results: The challenge of bridging the gap between ehalth research and health action. *Health Promotion International* 9: 289.

Dewatripont, M., Jewitt, I. and Tirole, J. 1999. The economics of career concerns, Part I: Comparing information structures. *Review of Economic Studies* 66: 183-198.

Gates Foundation. 2019. *Evaluation Policy.* Bill and Melinda Gates Foundation. Accessed online, February 19, 2019. https://www.gatesfoundation.org/How-We-Work/General-Information/Evaluat Policy

Gentzkow, Matthew and Emir Kamenica. 2017. Competition in Persuasion. *Review of Economic Studies* 84: 300-322.

Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings and Christel M.

J. Vermeersch. 2016. *Impact Evaluation in Practice, Second Edition*. Washington, DC: Inter-American Development Bank and World Bank.

Gustafsson-Wright, Emily, Sophie Gardiner and Vidya Putcha. 2015. *The potential and limitations onf impact bonds: Lessons from the first five years of experience worldwide*. Washington DC: Brookings Global Economy and Development Program.

Head, Brian W. 2010. Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society* 29: 77-94.

Hirshleifer, Jack. 1971. The Private and Social Value of Information and the Reward to Inventive Activity. *American Economic Review* 61: 561-574.

Holmstrom, B. 1999. Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies* 66: 169-182.

Hulko, Artem and Mark Whitmeyer. 2018. Information Provision in a Sequential Search Setting. Mimeo.

Karlan, Dean and Jacob Appel. 2011. *More than Good Intentions: How a New Economics is Helping to Solve Global Poverty*. New York: Dutton.

Li, Fei and Peter Norman. 2018. *Sequential Persuasion*. Mimeo.

MasterCard Foundation. 2015. *Research and Evaluation Policy*. http://www.mastercardfdn.org/learning/

Mebrahtu, Esther. 2002. Perceptions and practices of monitoring and evaluation: international NGO experiences in Ethiopia. Development in Practice 12: 501-517.

Merchant-Vega, Nina. 2011. Practical Challenges of Rigorous Impact Evaluation in International Governance of NGOs: Experiences and Lessons from the Asia Foundation. *The Asia Foundation*, Occational Paper No 8.

Morris, Stephen, and Hyun Song Shin. 2002. Social Value of Public Information. *American Economic Review* 92: 1521-1534.

Sappington, David. 1986. Commitment to Regulatory Bureaucracy. *Information Economics and Policy* 2: 243-258.

Sappington, David. 1991. Incentives in principal-agent relationships. *The Journal of Economic Perspectives* 5: 45-66.

Taylor, Curtis, and Yildirim, Huseyin. 2011. Subjective Performance and the Value of Blind Evaluation. *Review of Economic Studies* 78: 762-794.

van Teijlingen, Edwin R. and Vanora Hundley. 2001. The importance of pilot studies. University of Surrey *Social Research Update*. Issue 35.

6 Proofs

Proof of Proposition 1. The decision maker will implement the measure s proposals with the highest q_i. Define Q as the threshold value such that those with $q_i > Q$ receive funding, those with $q_i < Q$ do not. We make the following observations about agent strategies. First, agents would prefer the lowest q_i such that they receive funding. Thus, there must be a mass point on $q = Q$ and no investment in q above this mass point. Second, there if $Q < 1$, then the share of agents with $q = Q$ must equal s. To see this, not that probability of funding when $q = Q$ must equal 1. Otherwise, agents would have an incentive to deviate from $q = Q$ to marginally increase their q and guarantee funding. This means the share cannot exceed s. Furthermore, agents must not have an incentive to marginally decrease q, which means that there must not be capacity to fund those with lower q and the share at Q must not be less than s. Third, if $Q = 1$, then the share of agents with $q = Q$ must be at least s. Otherwise, agents would have an incentive to reduce their investments. Forth, agents strictly prefer $q_i = 0$ to any other $q_i < Q$. Since no $q_i < Q$ receives funding, it is better to choose the least costly q.

The above results imply that each agent chooses either $q_i = 0$ or $q_i = Q$ for some common $Q \in (0, 1]$. The costs of investment are strictly decreasing in agent ability (ρ_i). Therefore, there exists a threshold $\bar{\rho}$ such that agents with $\rho_i < \bar{\rho}$ strictly prefer 0, those with $\rho_i > \bar{\rho}$ strictly prefer Q and those with $\rho_i = \bar{\rho}$ are indifferent. Suppose that we are in the case where $Q < 1$. Here, an agent receives $1 - Q^2/(2\rho_i)$ from choosing $q_i = Q$, and receives 0 from $q_i = 0$. $\bar{\rho}$ is the value of ρ_i such that these payoffs are equal. Furthermore, we know from above that share s of the population must have $\rho_i > \bar{\rho}$. Thus, $\bar{\rho} = 1 - s$. This means,

$$1 - \frac{Q^2}{2(1-s)} = 0 \to Q = \sqrt{2(1-s)}.$$

Given that $q_i \in [0, 1]$, it must be that $Q \in [0, 1]$. This is the case when $s \geq 1/2$. This establishes the equilibrium for the case where $s > 1/2$.

For the case where $s \leq 1/2$, Q is at its upper limit of 1. In this case, an agent with $q_i = 1$ receives funding with probability equal to ratio of available funding over those who choose the maximum investment: $s/(1 - \bar{\rho})$. Thus, $\bar{\rho}$ solves $s/(1 - \bar{\rho}) - 1^2/(2\bar{\rho}) = 0$, or $\bar{\rho} = 1/(1 + 2s)$. This describes the equilibrium for the case where $s \leq 1/2$.

Proof of Corollary 1. When $s \leq 1/2$, at least s agents invest the maximum in project quality, guaranteeing high value. Therefore, the decision maker implements s high value projects. When $s > 1/2$, the decision maker awards funding to the s projects that have expected quality $\sqrt{2(1-s)}$, making the expected measure of high value projects equal to $s\sqrt{2(1-s)}$, which equals $1/2$ when $s = 1/2$ and achieves its maximum at $s = 2/3$.

Proof of Proposition 2. Let us start with the case in which there is rationing among high-type projects. Conditional on being $\tau_i = H$, define $\phi < 1$ to be the probability of being accepted. Each agent's optimal quality choice maximizes expected payoff $\phi q_i - q_i^2/(2\rho_i)$, and therefore $q_i = \phi\rho_i$. Given this, the overall measure of high-value projects developed is

$$\int_0^1 \phi\rho_i f(\rho_i)\, d\rho_i = \phi/2.$$

For consistency ϕ must be the share of accepted projects among those that are high value. Thus, $\phi = s/(\phi/2)$, which implies $\phi = \sqrt{2s}$. And, to be consistent with rationing, it must be that $s \leq 1/2$. Hence, actual measure of good projects in this case is $\phi/2 = \sqrt{s/2}$. Since only measure s are accepted, the decision-maker's payoff equals s.

Each agent's investment in quality is calculated as $q_i = \phi\rho_i = \rho_i\sqrt{2s}$.

Now, suppose that the decision-maker accepts all projects with $\tau_i = H$ and a ψ fraction of the projects with $\tau_i = L$. Each agent's optimal quality choice maximizes expected payoff $q_i + (1 - q_i)\psi - q_i^2/(2\rho)$, and is therefore equal to $q_i = (1 - \psi)\rho_i$. Given this, the overall measure of good projects developed is

$$\int_0^1 (1 - \phi)\rho_i f(\rho_i)\, d\rho_i = (1 - \phi)/2.$$

and overall measure of bad projects developed is $1 - (1 - \phi)/2$. For consistency ϕ must be the share of projects accepted among those that are bad

$$\begin{aligned} \phi &= \frac{s - (1 - \phi)/2}{1 - (1 - \phi)/2} \rightarrow \\ \phi &= \sqrt{1 - 2(1 - x)} \end{aligned}$$

Since all projects with positive reports are accepted, the decision-maker payoff is equal to the measure of good projects:

$$(1 - \phi)/2 = \left(1 - \sqrt{1 - 4\mu(1 - x)}\right)/2.$$

Each agent's investment in quality is calculated as

$$q_i = \left[1 - \sqrt{1 - 4\mu(1 - x)}\right]\rho_i.$$

Proof of Proposition 4. Suppose $G < 1$ and no rationing at G. An agent with ρ_i has a payoff $q_i/G - q_i^2/(2\rho)$ from choosing $q_i < G$ and payoff $1 - G^2/(2\rho)$ from choosing $q_i = G$. The former case is equivalent to investing less, but making up for it with more-informative evaluations. The latter case is equivalent to choosing $\gamma = 0$ (i.e., an uninformative signal), but investing sufficiently in quality to get funding. Among the agents who choose $q_i < G$,

$$q_i = \rho_i/G \quad \text{and} \quad u_i = \frac{\rho_i/G}{G} - \frac{(\rho_i/G)^2}{2\rho_i} = \frac{\rho_i}{2G^2}.$$

Since $q_i < G$, it follows that $\rho_i < G^2$. Note it is easy to check that for $\rho_i < G^2$, investing ρ_i/G is preferred to deviating:

$$\begin{aligned} \frac{\rho_i}{2G^2} &> 1 - \frac{G^2}{2\rho_i} \\ \frac{\rho_i}{2G^2} - (1 - \frac{G^2}{2\rho_i}) &= \frac{1}{2G^2\rho_i}\left(\rho_i - G^2\right)^2 > 0 \end{aligned}$$

If $\rho_i > G^2$, then the optimal quality is G. Hence, all types with $\rho_i < G^2$ prefer to exert effort less than G and then give a more informative signal, whereas types with $\rho_i > G^2$ exert effort G and always claim to have a good outcome. Hence, mass of agents with projects that exactly hit threshold G is equal to

$$\int_0^{G^2} \frac{\left(\frac{\rho_i}{G}\right)}{G} d\rho_i + \left(1 - G^2\right).$$

Because no rationing is possible at G, unless G is 1, it must be that if there is no rationing in equilibrium, then this expression equals s. Solving for G gives $G = \sqrt{2(1-s)}$. For $G \leq 1$, it must be that $s \geq 1/2$.

If there is rationing, then $G = 1$, and thus, $s \leq 1/2$.

Proof of Proposition 4. Follows from a straightforward comparison of the payoffs previously derived.