Creative Professions

Canice Prendergast*

April 16, 2018

Abstract

This paper offers a dynamic model of motivating agents in creative settings. Creativity is modeled as opportunities to map a landscape, where agents sequentially choose where to map and how hard to work. The central assumption is that creativity is hard to motivate as the inputs required for solving the unknown are so different to those typically used. Despite efficiency entailing agents choosing the point furthest from known landmarks, we show how limited creative ambition can be a way of alleviating agency concerns. This outcome is consistent with a considerable body of existing empirical evidence. This implies dynamics of discovery where breakthroughs are followed by periods of local innovations. This pattern of discovery repeats over and over. We also show that in these intervals of limited ambition, agents who attempt the most creative projects are least well paid. This is the inverse of the efficient outcome. We also show how a field can permanently or temporarily halt before its efficient point.

*University of Chicago Booth School of Business. Email: canice.prendergast@chicagobooth.edu.
Many thanks to Wouter Dessein for helpful comments.

Creativity involves a leap into the unknown. Much has been written, mostly outside economics, on how to encourage such leaps. Here we offer a dynamic agency model of creativity to reflect that literature. The central assumption is that creativity is hard to motivate as the inputs required for solving the unknown are so different to those typically used. Our focus is on identifying circumstances when this results in creativity being limited. In these cases, creative fields develop by breakthroughs being followed by many attempts at more minor problems, despite efficiency mandating more creativity. This pattern repeats over and over, and during these intervals of limited ambition, the most creative individuals are the least well paid. As a result, creative professions - business innovation, the arts, academia, etc. - at times appear to be nibbling around the margins of prior findings, rewarding most those who work on less important problems, but do so as a way of inducing effort exertion.

The model relies on creativity being hard to motivate. Most of us know examples of extraordinary achievements that were initially treated with derision - Stravinsky, Andy Warhol, Robert Goddard[1], and so on - but there is also considerable academic research that we discuss below. As one example from Siler et al., 2015, Figure 1 plots the citations of all works submitted for publication to three top medical journals - *Annals of Internal Medicine, The British Medical Journal*, and *The Lancet* - from 2003 to 2004. Of 808 papers that were eventually published anywhere, *every one* of the 14 most cited papers was rejected by (at least one of) these outlets, and published elsewhere. Of these, 12 were desk rejected without review. We model such issues by assuming that, in the absence of clearly achieved success, agents are evaluated on the inputs used - mastery of the literature, technical complexity, visual familiarity, and so on. Yet these are harder to interpret if the distance between the attempted project and what is known is large. Allied to the inherent randomness of succeeding at creative endeavors, this implies that it becomes hard to motivate creative agents.

Creativity is modeled here as a demand for novelty. A consumer matches actions to the height of an unknown landscape, where expected losses derive from being uncertainty about the landscape. The landscape is governed by a Brownian motion process, where the consumer knows more about points close to known landmarks. As a result, his willingness to pay is greatest for information on points furthest from those landmarks.[2] Creative endeavors are then attempts to map points far from the known landmarks.

To get that information, firms hire agents over time, choosing both where to map and how

[1] The father of rocket propulsion.

[2] This could be a pharmaceutical company seeking a treatment for a disease about which least is known. Or a fashion or art setting where the consumer wants to be offered something that distinguishes him from current fashion.

hard to work doing so. Modeling creativity in this way not only allows us to model creative ambition at one point, but also to show how the landscape is filled in over time, reflecting its temporal development.[3] Our interest is in the interaction between creativity and agency concerns. To isolate this, we assume that, holding agent effort constant, the probability of success is independent of which point on the landscape is chosen. (We show below that the qualitative results are robust to relaxing this assumption.) Allied to the demand for novelty, this implies that in the absence of agency issues, projects are carried out in order of their creativity, with expected compensation convex in creativity.

Despite this, our focus is on how professions instead encourage marginal innovations, where those who take on smaller problems are often better compensated than their more creative counterparts. The reason is an agency problem, where, due to limited liability on the part of agents, firms cannot induce efficient effort by paying agents solely on their successes. (This agency distortion is for analytic simplicity.) As a result, the agent is additionally paid on another signal of performance, which we call inputs. These reflect signals which are related to performance, but are not performance itself. Academia and the arts offer good examples of such inputs: for instance, while it is hard to calibrate the impact of a piece of economic theory on the world, it is not so hard to track the difficulty of its theorems. Furthermore, we assume that these inputs are more familiar when the problem being tackled is closer to existing successes. For example, techniques used to solve problems close to known results will likely be quite similar.

Then consider a case where, with the optimal compensation contract and the most creative available task, the agent is not sufficiently motivated to induce efficient effort. The key issue then becomes the interaction between effort and creative ambition. On the positive side, more creative projects have higher returns, and so naturally induce more effort. However, there is a countervailing incentive, in that if success is not achieved, the inputs used to solve the problem are less familiar when carrying out a project far from one's neighbors.[4] This adversely affects the incentive to exert effort. When this latter effect dominates, professions have incentives to restrict ambition: researchers use the findings of close neighbors to bound

[3]This abstract setting for creativity allows us to reflect both commonly cited activities like the visual arts or music, but in addition areas like academia or inventions. See Manso, 2011, for another approach to this problem.

[4]This is likely as one's peers find these inputs hard to interpret. Professionals "receive some significant portion of their incentives from organized groups of fellow practitioners" (Wilson, 1989, p.60). Or as Hall, 1968, put it: "the belief that the person best qualified to judge the work of a professional is a fellow professional" (p.92).

their own activities, leaving open terrain unexplored. We characterize the circumstances under which this arises.

With such static agency concerns, we then address dynamics. The dynamics concern both whether the landscape is appropriately filled in over time, and the order in which this occurs. First, a field can become permanently "stuck", in the sense that if some agent succeeded at a project, it would open up profitable further opportunities for future agents, yet no agent is willing to attempt that first project. This cannot happen here without agency concerns. Second, even when a field does not get stuck in this way, we identify potential intervals where agents take successively less ambitious projects even though a more ambitious one remains available. We characterize the expected length of this interval. After the remaining opportunities become sufficiently small, agents revert to attempt a breakthrough again. However, once an agent succeeds, history repeats itself in the sense that the interval above occurs over and over. In this sense, the field is filled in the wrong temporal order. The dynamic analysis also has implications for pay: specifically, during the intervals above, the most creatively ambitious agent receives the lowest expected compensation. We also show circumstances where globally agents work hardest on the least important problems, and where the most creative are least well paid. These outcomes are the inverse of the efficient outcome without agency problems.

These agency issues also allow empirically relevant comparative statics. First, there is considerable empirical evidence that creativity is encouraged by the intrinsic motivation of agents. The model predicts such a link not through any direct assumed channel but rather by relaxing agency problems: as those who are intrinsically motivated contribute more effort for any contract, there is less need to restrict creativity. Second, much of the empirical evidence focuses on how truly novel work is often not quickly recognized. We show how limiting ambition is heightened in such settings. In these cases of success not being recognized, "professional success is indexed by insularity" (Shapin, 2005).

Throughout the paper, we are conscious of benchmarking our results to existing empirical evidence. We first describe evidence on obstacles that novelty faces in gaining acceptance (Fiske and Taylor, 1991, Monahan et al., 2000, Eidelman et al., 2009, Wang et al., 2017, Calcagno et al, 2012). However, the model offers a potentially non-monotonic relationship between creativity and its returns - very local innovations may be well done, but are of little interest, while very creative activities may be weakly executed. We document empirical evidence in academia showing such a relationship, at least over the short run (Mukherjee et al., 2015, Uzzi et al., 2013). Finally, we discuss the evidence on how the effectiveness of incentive provision interacts with creativity (Erat and Gneezy, 2015, Kachelmeier et al., 2008), most notably Balsmeier et al., 2017, on the role of independent boards for innovation.

Above, we described how a field can become permanently stuck. We also provide an extension where a field lies fallow for some period of time, only to come back to life later. Here we do not rely on the dichotomy between inputs and success, but instead is motivated by empirical evidence that creative successes often take time to be recognized as such. This has the straightforward implication that agents may be unwilling to tackle creative problems, as it may take too long for them to be recognized. More interestingly, this has the further implication that an agent may be reluctant to attempt a problem previously tried by another agent, for fear that that earlier attempt might later be recognized as a success. This issue gives rise to a temporary shutdown of innovation. Yet this only happens when the opportunity is sufficiently creative: agents will continue to attempt projects with little creativity. In this way, the model offers a different notion of a stop-start process of creative innovation.

There is little formal work in economics about agency issues related to creativity, despite considerable empirical evidence on the subject. The closest literature to this work is on dynamic innovation, where a small number of influential papers (Jovanovic and Rob, 1989, 1990, Garfagnini and Stulovici, 2015) have addressed R&D settings where radical breakthroughs may be followed by more marginal contributions. Most formally similar is Garfagnini and Stulovici, who also use a Brownian motion landscape analogy. This work differs from that literature in many ways. First, that literature generates outcomes from an assumption that incremental innovation is costless while radical innovation incurs a fixed positive cost. One of the primary interests of the paper is in using an agency lens to unwrap this, in a way that reflects the empirical literature. Second, that literature is concerned with (typically business) settings where - in our language above - firms locate their product choices where the landscape is at its highest perceived point. In that literature, firms can profitably copy their predecessors, by replicating their choice and getting their return. Here we address settings - academia, the arts, and business settings where predecessors have property rights - where novelty is necessary for profit, with no value for reproducing what has been done before. A more extensive discussion of these and other issues is provided below.

Section 1 describes a model of creativity. We then considers the agent's incentives to exert effort and the reason to limit ambition in Section 2. Section 3 considers the dynamics of the model. We discuss evidence in Section 5. We address some comparative statics in Section 4. Section 6 considers the extension that allows fields to become temporarily fallow.

Section 7 provides a series of other extensions. Among these is showing how our results qualitatively continue to hold when more creative project have lower success rates for any level of effort. Section 8 concludes.

1 A Model of Creativity

There is a landscape or terrain defined over a normalized support $[0, 1]$, and at discrete point i, the terrain has height h_i. The terrain is linked across locations, moving according to a Brownian Motion with no drift, $dh_i = \sigma_h h_i dW_i$, where W_i is a Weiner process. As the discrete increments i become small, $h_b - h_a \sim N(0, \mid b - a \mid \sigma_h^2)$, where σ_h^2 is the variance of the process.

The landscape generates surplus in periods $t = 0, 1, 2, 3, ...$ When the game begins in period $t = 0$, the height of the terrain is known at the two end points (h_0 and h_1), but is unknown in the interior.[5] Without additional information, $h_i \sim N(ih_0 + (1-i)h_1, i(1-i)\sigma_h^2)$, where i close to 0 or 1 is well known, but least is known for $i = \frac{1}{2}$.

In each period t of the game, there are three players: a consumer, a firm, and an agent.

Consumers: A short run (one period) consumer derives welfare from the landscape in period t.[6] "Consumer t" does so by choosing a policy H_{it} for each location i. His expected payoff depends on matching the investment to the state of nature and is given by a quadratic loss function

$$L_t(h_i, H_{it}) = -\int_0^1 (h_i - H_{it})^2 di. \tag{1}$$

As one example, the consumer could be a hospital or pharmaceutical company seeking a cure H_{it} to a disease with characteristics h_i, where some diseases are close to existing knowledge and others less close.[7] Unless he purchases information, consumer t is uninformed about h_i beyond the prior distribution.

The consumer can purchase information from a firm. Consider the value of information to consumer t. Let h_{it} be the expected value of h_i after any information has been collected in period t. Consumer t chooses H_{it}^* to minimize $-\int_0^1 (h_{it} - H_{it})^2 di$, and sets $h_{it} = H_{it}^*$. As the

[5] This is for simplicity. We show below that the results are robust to relaxing the assumption of known end points.

[6] A single consumer is for simplicity.

[7] There are also mappings to cultural innovation, where an art collector may be more informed about what constitutes good painting (or a particular kind of painting), but has less information on something like performance art, and his returns to information may be higher as a result.

prior is h_i, the consumer's expected utility without additional information is the Brownian bridge $V_0 = -\int_0^1 i(1-i)\sigma_h^2 di = -\frac{\sigma_h^2}{6}$. Then consider the value of identifying some set of points by period t, characterized by a Bayesian belief of the variance of h_i given by σ_{it}^2. The consumer's expected utility with $h_{it} = H_{it}^*$ is then

$$V_t = -\int_0^1 \sigma_{it}^2 di, \tag{2}$$

which implies that the consumer t is willing to pay up to $V_t - V_0 = -\int_0^1 \sigma_{it}^2 di + \frac{\sigma_h^2}{6}$ to receive information on h_{it}.

The Agent: Information on the landscape is provided by the effort of agents, hired by a competitive market of firms. In each period, a single agent can potentially be hired by a firm (we call the agent hired in period t "agent t") to map a point on the landscape. As in Jovanovic and Rob, 1990, we assume that each agent lives for only one period, and that their order is exogenous. (We discuss this assumption below.) The agent is assumed to maximize wages minus effort costs and has a reservation utility $\overline{U}$.

The agent attempts to identify a single h_j, $j \in [0,1]$. Success in period t, which we denote Λ_t, is discrete. If success has occurred, $\Lambda_t = 1$, and h_j is revealed; otherwise, $\Lambda_t = 0$, and nothing about h_j is revealed.[8] The agent incurs unobserved effort at cost e_t. Let $\pi(e_t) \geq 0$ be the probability of success, where $\pi' > 0$, $\pi'' < 0$, and $\pi(\infty) < 1$. In order to isolate agency issues, this is assumed to be independent of location.

Firms: There is a competitive market of firms. In each period, a firm can employ an agent to identify a point on the landscape, and can sell that information to the consumer. Firms are infinitely lived, maximize profits of price minus wage costs, and discount at a rate δ. Consumer t receives a take-it-or-leave-it offer from the firm who employed agent t, where h_{it} is revealed for a price p_t. (As will be clear below, it is incentive compatible for no other firms to make offers to the consumer.)

We are interested in settings where there is a demand for novelty. To do this, we assume that there is no value to mimicking a predecessor. We do so by giving the earliest innovator property rights on her discoveries. To model this, we assume that a firm receive profit not

[8]The reason for the stark distinction between success and failure is to allow clean determination of dynamics. In particular, if success is observed with noise, some projects will partially succeed. In that case, the next agent will have to make a decision between finishing off the job on the old location or striking out elsewhere. The discreteness here allows us to avoid this.

just from its current consumer but also from all future consumers. Specifically, of the price p_t paid by consumer t, the firm that employed an agent in a previous period j receives $V_j - V_{j-1}$ from this. As the firm makes a take-it-or-leave-it offer to the consumer, it charges $p_t^* = V_t - V_0$.[9] As this is equivalent to $p_t^* = \sum_{j=1}^{t}(V_j - V_{j-1})$, firms are therefore only rewarded for incremental discoveries, receiving a net present value of $\frac{V_t - V_{t-1}}{1-\delta}$ from a period t discovery.[10]

It is useful to simplify the willingness to pay of the consumer. For some point i, let r_t be the distance between the nearest two known neighbors in period t. We call larger r_t more "creative ambition". We show in the Appendix that surplus is maximized by mapping the midpoint of the partition, with surplus from success of

$$V_t - V_{t-1} = \frac{\sigma_h^2 r_t^2}{12}. \tag{3}$$

If no success is achieved, then $V_t = V_{t-1}$. The price paid by the consumer is then $p_t^* = \frac{\sigma_h^2 \sum_{j=1}^{t} r_j^2 \Lambda_j}{12}$. As this price is charged to all future consumers, the expected return to the firm upon success (before wage costs) is

$$S(r_t) = \frac{\sigma_h^2 r_t^2}{12(1-\delta)}. \tag{4}$$

We assume that firms can choose r_t. Professions can often directly affect the ambition of its members by choosing what to fund, publish, or exhibit. As a result, we assume r_t is contractible. (We show below that the same results arise if payoffs can be contracted upon.)

Information: There are two signals on the agent's effort. First, Λ_t is observed. The agent is also rewarded on another signal of performance that we call inputs, reflecting pieces of a puzzle that may lead to success.[11] These input signals are valuable as paying solely on success does not lead to efficient effort below. For now, we treat the inputs as simply additional information on e but below we formalize the link between inputs and success. We denote outcomes on the inputs by $I_t \in \{0, 1\}$, where we assume that the probability of $I_t = 1$

[9] In some creative settings, firms can sell goods to consumers to capture these rents; for example, galleries selling art, studios selling movies, businesses capturing monetary returns to innovation, and so on. In others, it is the status associated with being associated with success, such as museums hosting breakthrough shows, or universities whose employees win major awards.

[10] Note that this price apportionment satisfies budget balance. Also note that given this apportionment of rents, no other firms benefit from making an offer to the consumer.

[11] Examples of such measures could be the technical complexity of a scientist's work, the narrative structure of a poem, the understanding of the previous literature for a social scientist or artist, formal rigor, and so on.

is given by $\Gamma(e_t, r_t)$. This is revealed at the same time as Λ_t. We assume that $\Gamma_e > 0$, and $\Gamma_{ee} \leq 0$. We discuss the relation to r below. We assume for simplicity that, conditional on effort, π and Γ are independent.

The paper is largely concerned with how creativity is affected by agency problems. Here we use a simple tractable source of such agency costs, namely limited liability for agents. Specifically, let w_t be the wage paid to the agent:

Assumption 1: $w_t \geq \overline{w} > 0$, where $\overline{w} \leq \overline{U}$.

The second part of this is simply that limited liability arises at a point no higher than their market wage.

The timing of the model is as follows. Entering period $t = 1, 2, 3, ...$, all firms observe $\sigma^2_{i(t-1)}$, and agent t and consumer t are born. Firms then simultaneously offer a contract (w_t, r_t) to agent t, where the wage can be conditioned on Λ_t and I_t. The agent accepts at most one contract, exerts effort e_t, which results in a realization of success Λ_t and input performance Γ_t. Consumer t is offered h_{it} for a price p_t by the firm employing the agent. If the consumer accepts, prior firms are paid according to $V_j - V_{j-1}$. If the consumer rejects, no transfers are made. Then the agent is paid $w_t \geq \overline{w}$, the consumer choose H_{it}, σ^2_{it} is computed, the agent and consumer die, and the next period begins.

Benchmark without Agency Concerns As a benchmark, consider the outcome where effort is contractible. This maximizes myopic expected surplus.[12] If the participation constraint is slack, the choice of effort maximizes $\pi(e_t)S(r_t) - e_t$, and its optimal choice $e^{**}_t(r_t)$ is given by

$$\pi'(e^{**}_t)S(r_t) = 1. \tag{5}$$

As returns are increasing in r in (4) and π is independent of r, the agent should work on the largest remaining partition, which has size r^{**}_t. She exerts effort $e^{**}_t(r^{**}_t)$, and is paid $w_t = \pi(e^{**}_t(r^{**}))S(r^{**}_t)$. Furthermore, effort is increasing in (a) ambition (r^{**}_t), (b) lack of knowledge of the landscape (σ^2_h), and (c) patience. If $\pi(e^{**}_t(r^{**}_t))S(r^{**}_t) - e^{**}_t \geq \overline{U}$, a contract with these features is accepted, and otherwise discovery ends.

[12] This can be seen by backward induction. The final agent who attempts a project receives his expected one period surplus, as there is no future surplus. Given this, she chooses the outcome that maximizes expected myopic surplus. But if this occurs for the final project attempted, then the second to final agent does likewise, as she extracts no future surplus. Unraveling in this way for all agents implies that all choices maximize myopic surplus.

This outcome has the intuitive feature that the most ambitious project is tackled first: the field evolves with a project of size $\frac{1}{2}$, two of $\frac{1}{4}$, four of $\frac{1}{8}$, and so on. Note also that the probability of success declines over time, as agents work hardest on the more creative projects done first. Finally, expected compensation of agents declines in a convex manner over time, as larger partitions disappear. This process occurs until the largest remaining partition does not offer the agent at least expected utility of $\overline{U}$, at which point exploration stops.

Observations Before addressing the agency problem, it is worth making some observations. First, the outcomes of the model rely on two features: (i) that goods are ordered by similarity to existing entities, and (ii) that consumers have greatest demand for those goods least similar to those existing goods. We have formalized this through mapping a landscape. Alternatively, we could have simply assumed such preferences, and the results below hold. As one such example, a customer could have a demand for fashion or art, and where their demands are focused on avoiding the status quo. Second, note that there is a single agent in each period. This allows us to avoid inefficiencies in entry in research tournaments (Taylor, 1995, McAfee and Fullerton, 1999), or complications that arise when points can be simultaneously discovered by more than one agent.[13]

Third, giving firms property rights avoids another reason for lack of novelty, namely that a firm might want to mimic a predecessor. Here mimicking offers no surplus. Related to this, consumers are assumed to be uninformed about h_i beyond the prior unless they purchase information. This is simply so that firms also have property rights over consumers, such that they cannot free ride on previous discoveries. An alternative assumption would be to allow them to observe $h_{i(t-1)}$ but require them to pay previous innovators for using it.[14]

Finally, note that the model does not imply that the consumer observes payoffs in each state - instead, the loss function is an expected payoff. For example, it could imply a setting where randomly some state i is realized, or one where only coarse realizations of outcomes arise. In the hospital setting, a patient could get a particular illness or not, and the outcome could be coarse (patient survives/does not survive). For this reason, we assume that the

[13] From a more positive perspective, our interest is partly in how projects can be done in the wrong order, with less important problems being done before more important ones. Having a limited supply of agents allows us to model this.

[14] One alternative might be where property rights do not cover consumers, where consumer t gets to use $h_{i(t-1)}$ for free. Then he is only willing to pay for the current update. Then the value of an innovation is $S(r)(1-\delta)$ rather than $S(r)$. This has the same features that arise below, but would reduce effort for a different reason to below.

consumer cannot offer useful information to the discovery process.

2 The Agency Problem

The myopically efficient benchmark abstracts from the need to encourage effort. We now address this.

2.1 The Optimal Contract

Agent compensation is potentially based both on success and performance on the inputs, $w_t(\Lambda_t, I_t)$. Consider the choice of contract (w_t, r_t), where the set of available partitions inherited from the last period are of size $\mathbf{r_t}$. As there is competition between firms, the period t contract maximizes the utility of agent t. This involves maximizing myopic surplus each period.[15] Let $\hat{U}(e^*(r_t, w_t), r_t)$ be the agent's expected utility, where

$$\hat{U}(e^*, r_t) = S(r_t)\pi(e^*) - e^*, \tag{6}$$

and $e^*(r_t, w_t)$ is equilibrium effort for a project of size r_t and contract w_t. The firm chooses r_t and w_t to maximize (6), subject to $e^*(r_t, w_t)$, $\hat{U}(e^*, r_t) \geq \overline{U}$, and $r_t \in \mathbf{r_t}$.[16]

To reduce the dimensionality of the problem, begin by considering e^* with the optimal contract $w_t^*(\Lambda_t, I_t)$ for a given r_t. This is straightforward. As the agent is risk neutral above $\overline{w}$, the optimal incentive contract pays the agent only in the state that is most informative of effort, namely when both $\Lambda_t = 1$ and $\Gamma_t = 1$. Hence, $w_t^*(1,1) = w_{\Lambda I}$; otherwise $w_t^*(\Lambda, I) = w_0 \geq \overline{w}$. Conditional on r, it is independent of t. (From now on, we ignore time subscripts unless necessary.) The agent then chooses effort to maximize $\pi(e)\Gamma(e,r)w_{\Lambda I} + (1 - \pi(e)\Gamma(e,r))w_0 - e$. For notational convenience, let $\frac{\Gamma_e(e,r)}{\Gamma(e,r)} \equiv g(e,r)$. Effort e^* is then given by

$$\left(\pi'(e^*) + \pi(e^*)g(e^*, r)\right)(w_{\Lambda I} - w_0) = 1. \tag{7}$$

[15] This can be seen by backward induction. Let $\overline{r}$ be the final sized partition attempted by agents - beyond that the agent's expected utility is below $\overline{U}$. For that partition, the objective of the agent is to maximize its myopic return $\pi(e)S(\overline{r}) - e(\overline{r})$, as there is no future return. But if the final agent does this, any agent attempting the next to last sized partition will do likewise. Unraveling then implies that so also do all earlier agents. Said another way, a firm could conceivably want to choose a low expected return project today because it could increase payoffs in the future. However, there is no way to compensate today's agent for doing so. As a result, competition implies that the firm needs to make non-negative profits in each period.

[16] There is a firm solvency constraint that we are ignoring here. Specifically if π is sufficiently low, the firm may not break even if it has to pay all failed workers $\overline{w}$. This issue is far from our interests here, and is ignored.

The choice of $w_{\Lambda I}$ and w_0 is constrained by budget balance,[17]

$$\pi(e^*)S(r) \leq \pi(e^*)\Gamma(e^*,r)w_{\Lambda I} + (1-\pi(e^*)\Gamma(e^*,r))w_0. \tag{8}$$

As a result, the simplified program chooses $r, w_{\Lambda I}$, and w_0 to maximize (6), subject to (7), (8), $\hat{U}(e^*,r) \geq \overline{U}$, and $r \in \mathbf{r}$.

To see the limits of compensation contracting, consider the wage contact that offers the greatest incentives, where $w_0 = \overline{w}$ and $w_{\Lambda I}$ is the maximum payment consistent with budget balance, given by $w^*_{\Lambda I} = \frac{S(r)}{\Gamma(e^*,r)} - \frac{(1-\pi(e^*)\Gamma(e^*,r))\overline{w}}{\pi(e^*)\Gamma(e^*,r)}$. Then the maximum feasible effort for a project r is given by $(\pi'(e^*) + \pi(e^*)g(e^*,r))\,(w^*_{\Lambda I} - \overline{w}) = 1$, or

$$(\pi'(e^*(w^*_{\Lambda I})) + \pi(e^*(w^*_{\Lambda I}))g(e^*(w^*_{\Lambda I}),r))\left(S(r) - \frac{\overline{w}}{\pi(e^*(w^*_{\Lambda I}))}\right) = 1, \tag{9}$$

where the second order condition is assumed to hold.[18]

Now consider the impact of changing r with that contract. It is useful to begin by addressing how a marginal increase in r affects the agent's utility. (The firm does not make such local choices - instead it chooses between the midpoints of a discrete set of available partitions - but the logic is useful.) Note that

$$\frac{d\hat{U}(e^*,r)}{dr} = \pi(e^*)S'(r) + \hat{U}_e\frac{de^*}{dr}, \tag{10}$$

where $\hat{U}_e = 1 - (\pi' + \pi g)\left(S(r) - \frac{\overline{w}}{\pi}\right)$ reflects the agency problem, with higher effort valued at its marginal return $\hat{U}_e$. The relevant issue here is whether the agency problem can be resolved only through compensation. To understand $\hat{U}_e$, it is useful to rewrite (9) as

$$[\pi'(e^*)S(r) - \frac{\pi'(e^*)\overline{w}}{\pi(e^*)}] + g(e^*,r)[\pi(e^*)S(r) - \overline{w}] = 1. \tag{11}$$

The first bracketed term identifies maximum possible effort with a contract was based only on success ($w_t(1,1) = w_t(1,0)$). Limited liability always causes effort to fall below its desired level: without further signals, efficiency ($\pi'(e^*)S(r) = 1$) can only be attained if $\overline{w} = 0$. The second term allows additional information on inputs to mitigate (and potentially overturn) the agency problem.[19] Our interest is in cases where agency issues cannot be resolved by compensation alone. This arises when the agency problem (here parameterized by $\overline{w}$) is sufficiently important, as given by Assumption 2.[20]

[17]As the firm does not observe effort, it is based on an expectation e^*. However, that expectation is correct in equilbrium.

[18]This is given by $\left(\pi'' + \pi'(e^*)g(e^*,r) + \pi(e^*)\frac{\Gamma_{ee}\Gamma - \Gamma_e^2}{\Gamma^2}\right)\left(S(r) - \frac{\overline{w}}{\pi(e^*)}\right) < 0$.

[19]If there was no cost to using pay for performance, the firm could trivially replicate the efficient outcome above by setting incentives in any combination such that $w_\Lambda + w_I\frac{\Gamma'(e^{**},r^{**})}{\pi'(e^{**})} = S(r^{**})$.

[20]To show how limited creativity arises as a potential solution to agency concerns, we only need to make

Assumption 2: $U_e(e^*(w^*_{\Lambda I}), r) > 0$ or $\overline{w} > \pi(e^*(w^*_{\Lambda I}))\left(S(r) - \frac{1}{\pi'(e^*(w^*_{\Lambda I})) + \pi(e^*(w^*_{\Lambda I})g(e^*(w^*_{\Lambda I}, r)}\right)$.

Assumption 2 implies that because effort is suboptimal, $w_{\Lambda I}$ will be set at its maximum level. To simplify notation, we do not condition e^* on this wage below.

The motivation for the paper is that creativity is hard to interpret and reward. We model this by assuming that the inputs used for local problems are more easily "interpretable" than are those used for problems from from known problems, via Assumption 3.

Assumption 3: $g_r < 0$.

Efficiency in agency setting is often determined by likelihood ratios, identifying the informativeness of signals. In that vein, g is the marginal return of effort, normalized by its average level. This assumption implies that it falls as agents attempt more creative projects. Note that $g_r = \frac{\Gamma_{er}\Gamma - \Gamma_e\Gamma_r}{\Gamma^2}$. One natural interpretation of this is that creativity per se does not affect average performance ($\Gamma_r = 0$) but that as performance measures become noisier ($\Gamma_{er} < 0$). We offer two examples of this in Section 7.

2.2 Single Period Creativity

The agent prefers a marginally less ambitious project if (10) is negative. As $S'(r) > 0$ and $\hat{U}_e > 0$ from Assumption 2, this requires that $\frac{de^*}{dr} < 0$. But

$$\frac{de^*}{dr} = \frac{(\pi' + \pi g)S'(r) + g_r\left(S(r) - \frac{\overline{w}}{\pi}\right)}{H}, \tag{12}$$

where H is an amended version of the second order condition above, and assumed to be positive.[21] The first term in the numerator of (12) is positive, because the surplus generated by more effort is increasing in creative ambition. Yet, from Assumption 1, the second term is negative. Substituting (12) into (10) yields

$$\frac{d\hat{U}(e,r)}{dr} = S'\left(\pi + \frac{\pi' + \pi g}{H}\right) + \frac{g_r\left(\pi S - \overline{w}\right)}{H}. \tag{13}$$

This provides the foundation for understanding the influences that arise below. First - and most obviously - the value of success is increasing in r. Second, there is a complementarity

Assumption 2 evaluated at the efficient project choice. We make this more extended assumption solely so we can restrict attention to contracts which have the maximum bonus payments.

[21] The term $H = -(\pi'' + \pi' g + \pi g_e)\left(S(r) - \frac{\overline{w}}{\pi(e^*)}\right) + \pi'(\pi' + g)\frac{\overline{w}}{\pi^2} > 0$ is amended from above because r is chosen before effort, and as a result, the optimal contract adapts.

between the agent's effort and the size of the prize ($\frac{d^2S}{dedr} > 0$). Both of these influences lead to taking a bite at the most ambitions (or creative) remaining partition. However, as $g_r < 0$, a more ambitious project may induce less effort as inputs are less informative. Because of this last effect, the effect of a more ambitious project on the agent's welfare is ambiguous. This illustrates the tradeoff. Below we consider when these agency issues are most likely to arise.

The agent does not make marginal partition choices as above. Instead, discrete options are available, but the logic is identical. Remember that partitions are optimally split in half. This implies that the agent potentially chooses between the largest remaining partition, of size $\frac{1}{2^k}$, and others of size $\frac{1}{2^{k+s}}$, where $k \geq 2$ and $s \subset \{1, 2, 3...\}$. Surplus from succeeding at the larger partition $\frac{1}{2^k}$ over $\frac{1}{2^{k+s}}$ is $\frac{S(\frac{1}{2^k})}{S(\frac{1}{2^{k+s}})} = \frac{1}{s^2}$. As a result, the agent then prefers the partition $\frac{1}{2^k}$ over $\frac{1}{2^{k+s}}$ iff

$$S(\frac{1}{2^k})[\pi(e^*(\frac{1}{2^k})) - \frac{\pi(e^*(\frac{1}{2^{k+s}}))}{s^2}] \geq e^*(\frac{1}{2^k}) - e^*(\frac{1}{2^{k+s}}). \quad (14)$$

More generally, the agent carries out this comparison for all available $k + s$ partitions and only prefers the largest one if (14) holds for all available k.

3 Dynamics of Creativity

Up to now, we have considered the static incentives of a single agent. How does the discovery process evolve over time? The simplicity of (14) allows us to characterize dynamics. Note first that the tradeoff above is not relevant until there are two successful agents (as there is no smaller partition than the most creative one available). Furthermore, there may be some later agents who have no choice.[22] Hence the results below arise in cases where an agent has a choice between a partition $\frac{1}{2^k}$ and others of size $\frac{1}{2^{k+s}}$. A little bit of notation is useful here. Let $U(k) \equiv \hat{U}(e^*(\frac{1}{2^k}), \frac{1}{2^k})$. Furthermore, let $\overline{k}$ be the smallest k such that $U(\overline{k} + 1) < \overline{U}$: hence, exploration stops after all partitions of size $\frac{1}{2^{\overline{k}}}$ are completed.

Consider an agent who is faced with the largest partition of size $\frac{1}{2^k}$. If that is her only choice, and $k < \overline{k}$, then she carries out that project. Of more interest is if there is a smaller

[22]On the first issue, the first agent chooses point $\frac{1}{2}$ so $r = 1$, and the agent after the first success chooses point $\frac{1}{4}$ or $\frac{3}{4}$ so $r = \frac{1}{2}$. However, the agent who moves after the second successful agent can choose a range of either $r = \frac{1}{8}$ or $r = \frac{1}{4}$. On the second, for example the agent who moves after three successes at points $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{3}{4}$ only has projects of size $\frac{1}{8}$ remaining.

partition. Define $\kappa(k)$ as the first value $k+1, k+2, ...,$ such that $U(k) \geq U(\kappa)$, or

$$S(r)[\pi(e^*(\frac{1}{2^k})) - \frac{1}{(\kappa-k)^2}\pi(e^*(\frac{1}{2^\kappa}))] \geq e^*(\frac{1}{2^k}) - e^*(\frac{1}{2^\kappa}). \quad (15)$$

In words, $\kappa(k)$ is the next smallest partition where the largest one k is preferred. As a benchmark for what follows, Proposition 1 immediately follows.

Proposition 1 *If $\kappa = k+1$ for all $k \leq \overline{k}$, then the agent always attempts the most ambitious project until project $\overline{k}$, at which point there is no further searching.*

Getting Stuck Before describing cases where projects are taken in a different order to the efficient benchmark, consider one other distortion. If effort is contractible, projects are done in order of their surplus, and if one agent chooses not to carry out the best remaining project, there is no smaller project worth exploring. This is not necessarily true here. This arises if $U(k) > \overline{U}$, for some $k > \overline{k}$. This can happen only if effort is higher for the less ambitious project. In words, if someone succeeded at a project of size $\frac{1}{2^k}$, this would open up smaller projects that future agents would attempt, but no one is willing to attempt the size k project. In this sense, a field can "get stuck" permanently.[23]

Now consider the case where $\kappa \neq k+1$ for some k, but where the field is not stuck, as $U(k) \geq \overline{U}$. Here the agent prefers to take a less ambitious project than the biggest one available, but will attempt the larger one if no project between $k+1$ and $\kappa - 1$ is available. Proposition 2 formalizes the the length of intervals where limited ambition arises, the sense of history repeating itself, and the returns to creativity.

Proposition 2 *Consider the first agent who is faced with a largest partition of size $\frac{1}{2^k}$ and at least one smaller option. Then if $\kappa \neq k+1$, there is an* interval *where agents do not attempt the most ambitious project. During this interval,*

- *The number of successful agents who pass on projects of size $\frac{1}{2^k}$ is $2^{\kappa-k} - 2$,*
- *The expected number of agents who attempt a project smaller than $\frac{1}{2^k}$ is $\sum_{i=1}^{\kappa-k-1} \frac{2^i}{\pi(e^*(\frac{1}{2^{k+i}}))}$.*
- *During the interval, the agents who attempt the most creative project are least well paid in expected terms.*

[23]The fact that there is a smaller project than $\frac{1}{2^k}$ that has welfare above $\overline{U}$ does not of course mean it is welfare improving for the larger project to be done. This requires a comparison of the welfare loss from the intervening projects, and in addition any later projects that might be done as a result.

- *After the interval ends, agents again attempt a project of size $\frac{1}{2^k}$. When success occurs, the interval above occurs again. This process of success followed by limited ambition arises a further 2^{k-1} times in total.*

After this exercise completes all projects of size $\frac{1}{2^{\kappa-1}}$, the agent chooses whether to explore a project of size $\frac{1}{2^\kappa}$, and the model continues.

The intuition for the interval duration is simply that if an agent prefers projects up to size $\frac{1}{2^{\kappa-1}}$ over one of size $\frac{1}{2^k}$, there are $2^{\kappa-k}-2$ possible partitions of sizes $\frac{1}{2^{k+1}}$ to $\frac{1}{2^{\kappa-1}}$. Furthermore, one of those smaller partitions is always available, so all will be completed before one of size $\frac{1}{2^\kappa}$ is attempted.[24] Note that we do not know the order of these $2^{\kappa-k}-2$ projects, as we cannot rule out further cycles of even more limited ambition within that interval.[25] This addresses the number of successes, not the number of agents. The probability of a success for a project of size $\frac{1}{2^{k+i}}$ is given by $\pi(e^*(\frac{1}{2^{k+i}}))$. As the probability of success is independent across partitions and time, and there are 2^{j-1} partitions in total of size $\frac{1}{2^j}$, the expected number of agents exhibiting limited ambition is $\sum_{i=1}^{\kappa-k-1} \frac{2^i}{\pi(e^*(\frac{1}{2^{k+i}}))}$.

This describes the dynamics that arise only until the *next* project of size $\frac{1}{2^k}$ is attempted. After success has occurred at the second project of size $\frac{1}{2^k}$, the exact same cycle outlined above occurs again, with identical expected outcome. There are 2^{k-1} partitions in total of size $\frac{1}{2^k}$, so the interval above occurs that number of times before all projects of size $\frac{1}{2^{\kappa-1}}$ are completed. In this sense, the outcome of the model is history repeating itself $2^{k-1}-1$ times.

Finally, consider pay. If effort could be contracted upon, the most creative agent is paid best: indeed, expected pay is weakly convex in creativity. During the intervals outlined above, the opposite is true. Here the lowest utility is attained by the agent taking the $\frac{1}{2^k}$ project, the largest available one. Furthermore, as her effort is lowest for that project, this implies that expected compensation is lowest for the agent who takes the most creatively ambitious project.

[24] As an example, consider the agent who moves after the first two successes at points $\frac{1}{2}$ and $\frac{1}{4}$. The biggest available project is at $\frac{3}{4}$. Let $\kappa = 6$, so that partitions up to $\frac{1}{32}$ are preferred to one of size $\frac{1}{4}$. The third agent chooses a location $\frac{1}{8}$ instead of $\frac{3}{4}$. This continues until an agent is successful at that location. The next agent can now choose a partition of size $\frac{1}{16}$ or $\frac{1}{4}$. But as $\kappa = 6$, the smaller partition is chosen, and so on. There are $2^4 - 2 = 14$ such partitions.

[25] As an example, assume that the agent faces a largest project of size $\frac{1}{4}$, but $\kappa = 6$. After the first $\frac{1}{8}$ is solved, the next agent chooses $\frac{1}{16}$, and that is her only smaller choice. However, the next agent can choose $\frac{1}{8}$, $\frac{1}{16}$, or $\frac{1}{32}$. While she prefers $\frac{1}{16}$ and $\frac{1}{32}$ to $\frac{1}{8}$, we have said nothing about whether she prefers $\frac{1}{16}$ to $\frac{1}{32}$. If she does, then all partitions of $\frac{1}{16}$ are done first. But if $\frac{1}{32}$ is preferred to $\frac{1}{16}$, there is another inversion of creative ambition. In that case, the order in which the interval after the first $\frac{1}{8}$ is solved is $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{32}$, $\frac{1}{16}$, $\frac{1}{32}$ and $\frac{1}{32}$.

Effort, Creativity, and Wages The dynamics above point to the possibility of agents preferring less ambitious projects for *some* ranges of ambition. However, can it be the case that agents globally exert most effort on the least important attempted problems? Relatedly, can the most creative agents be globally least well paid, as their effort is commensurately lower?

To address this, consider an example where $\overline{k} = 2$ so that only projects of size $\frac{1}{2}$ and $\frac{1}{4}$ are attempted. The willingness to take on a project potentially depends on the returns to any other available projects, and this simplifies the options. Remember that the marginal return to effort with maximum incentives is given by $(\pi'(e) + \pi(e)g(e,r))\left(S(r) - \frac{\overline{w}}{\pi(e)}\right)$. Then if

$$\left(\pi'(\frac{1}{2}) + \pi(e)g(e,\frac{1}{2})\right)\left(S(\frac{1}{2}) - \frac{\overline{w}}{\pi(e)}\right) < \left(\pi'(\frac{1}{4}) + \pi(e)g(e,\frac{1}{4})\right)\left(S(\frac{1}{4}) - \frac{\overline{w}}{\pi(e)}\right), \quad (16)$$

the agent works hardest on the least creative taken project.

This example also also allows the possibility that those who take on the most creative projects are globally least well paid. This arises if expected compensation $\pi(e^*(\frac{1}{2}))S(\frac{1}{2}) < \pi(e^*(\frac{1}{4}))S(\frac{1}{4})$.

So far, we have said little about the speed at which the landscape is filled. If effort could be contracted upon, the expected time between success is decreasing in r. By contrast, when the agency issues here imply limited creative ambition, the opposite occurs: slow breakthroughs on important problems are followed by more rapid successes on smaller problems. The expected number of periods until success is achieved for a project of size r is $\frac{1}{\pi(e^*(r))}$. In the example above, if more effort is exerted only on less ambitious problems, then projects of size $\frac{1}{e}$ take an average of $\frac{1}{\pi(e^*(\frac{1}{2}))}$. When success on this is finally achieved, it is followed by more rapid local successes at a rate $\frac{1}{\pi(e^*(\frac{1}{4}))}$ for the two smaller projects. n this way, the model predicts the possibility of long periods between breakthroughs, followed by a rapid series of local successes.

4 Agency Comparative Statics

The dynamics above rely on agency issues being sufficiently severe. Yet we have not fully explored when this is likely to occur. Here we provide comparative statics on the single period problem. Note that there are two signals on performance. As a result, we carry out two conceptual exercises: (i) when the severity of the agency problem changes for both signals, and (ii) when the ability to interpret one of the two signals changes. We deal with each in turn.

Intrinsic Motivation and Creativity There is a large literature on how intrinsic motivation tends to lead to more creativity. Classic references here are Deci and Ryan, 1985, and Amabile, 1997. See Hennessy, 2001, for a review.[26] The usual conduit in that literature is how intrinsic motivation implies a curiosity that leads to more creative solutions. This model offers an alternative link, where intrinsic motivation leads to creativity by relaxing agency concerns.

Consider a setting where intrinsic motivation is modeled as an agent being capable of exerting some effort at no cost. (The cost could be negative over some range with no change in results.) As effort is modeled in cost units, this is represented by a probability $\pi(0) = \nu$ that can be produced at no cost, where higher ν reflects more intrinsic motivation. Otherwise the model is unchanged, where we continue to assume that Assumption 2 holds. The analog to (10) is now

$$\frac{d\hat{U}(e^*, r)}{dr} = (\pi(e^*) + \nu)S'(r) + \hat{U}_e \frac{de^*}{dr}. \tag{17}$$

To see the role of intrinsic motivation, note that $\frac{d^2\hat{U}(e^*,r)}{drd\nu} = S'(r) > 0$, so that more intrinsically motivated workers benefit more from greater creativity. In simple terms, an intrinsically motivated agent is less likely to limit creativity, as she has to give up more. Proposition 3 immediately follows.

Proposition 3 *Creative ambition is weakly increasing in intrinsic motivation, ν.*

Note that this outcome was not hard-wired as the willingness to exert effort intrinsically did not depend on creativity per se.

When Success is Hard to Identify Much of the motivation for the agency problem above is that agents are less likely to be recognized for novel contributions. Yet the model above only considers limited liability as an agency constraint. There are two variants of recognition for creativity. First, it may never be recognized ($\pi_r < 0$). We deal with this below. Second, it could be that while success can generate payoffs to the firm, the agent does not get those returns. This could be because success comes too late for the agent to benefit. To address this, assume that the agent is rewarded for success with probability z, while performance on inputs is always observed. With this addition, the agent's incentives

[26] Much of this empirical evidence uses ex post self-reports of task interest to measure intrinsic motivation and correlates these responses with various measures of performance. However, there is some work that seeks external measures of intrinsic motivation. For example, Reeve and Nix, 1997, measures intrinsic motivation through the facial expressions of experimental subjects, and notes correlation to performance.

with the contract above are now given by

$$z\left(\pi'(e^*)+\pi(e^*)g(e^*,r)\right)\left(w_{\Lambda I}-\overline{w}\right)=1. \tag{18}$$

This extension, in itself, does not change outcomes because the optimal bonus can simply be scaled up proportionately to maintain incentives: instead of offering a marginal reward of $\left(S(r)-\frac{\overline{w}}{\pi(e^*)}\right)$, the optimal contract offers $\left(\frac{S(r)}{z}-\frac{\overline{w}}{z\pi(e^*)}\right)$, and nothing changes. This is no longer true if there are limits to how much bonuses can be raised. For example, there could be fairness constraints across workers, or it could be that the firm might be tempted to renege if rewards for success are too great. A simple way to address this is to additionally assume that there is an exogenous maximum bonus b^{max} above $\overline{w}$ that can be given to an agent.

In this case, the informativeness of the signals potentially becomes critical for the maximum feasible effort. Specifically, the optimal contract changes, in that it may no longer be efficient to pay a bonus only when both signals are successful. For example, if recognized success in both is rare enough, offering b^{max} in that state may offer little incentive, and so it becomes efficient to offer some bonus in states other than when both outcomes are 1. Specifically, if the b^{max} constraint is sufficiently binding such that $\overline{w}+(\Gamma+(1-\Gamma)\pi)b^{max}<\pi(e^*)S(r)$, we show in the Appendix that the optimal contract is instead to offer b^{max} if *either* signal is positive. Incentives are then given by $(z\pi'+(1-z\pi)\Gamma_e)b^{max}=1$, and so

$$\frac{de^*}{dr}=\frac{(1-z\pi)\Gamma_{er}}{-(1-z\pi)\Gamma_{ee}-(1-\Gamma_e)z\pi''+z\pi'\Gamma_e}<0 \tag{19}$$

if the second order condition holds. As a result, incentives are always encouraged by reducing creativity unlike the ambiguous outcome above.

Remember, however, that this only arises if wage payments are constrained by b^{max}. From (18), that is more likely as z falls. Hence the inability to observe success (coupled with constraints on wage payments) makes the the value of restricting creative ambition to induce effort more desirable.[27]

[27] This does not necessarily guarantee that the firm restricts creativity, as there is still the direct cost of $\pi S'(r^{**})$ from (13). Specifically, then let $w_{\Gamma\cup I}$ be the wage paid if either signal is a success, and w_0 be the wage if both fail. Then from the firm's budget constraint, $w_0=\pi S(r)-(z+(1-z)\pi)b^{max}$, and $w_{\Gamma\cup I}=w_0+b^{max}$, and so expected pay still depends on r.

5 Evidence

Now consider the existing empirical evidence beyond that on intrinsic motivation above. The first issue is resistance to novelty, on which there is a large literature in psychology using experimental data. For example, Fiske and Taylor, 1991, illustrate how uncertainty can be an aversive state to people, and that as a result, creative solutions make subjects uncomfortable. A particularly elegant illustration of this is Monahan et al., 2000.[28] Also related is Eidelman et al., 2009, who show what they call "existence bias" where people incorrectly favor that which already exists.

The second source of evidence on a potential penalty for novelty is from academia. Above, we showed the Siler et al., 2015, data on the rejection of the 14 most subsequently cited papers submitted to three major medical journals. The most systematic evidence on novelty in academia is Wang et al., 2017, who study over 750,000 Web of Science articles that were published in 2001. Using a cross-citation metric to measure novelty, they show two striking results. First, there is delayed acceptance for novel research - depending on how novelty is defined, work that has little novelty is more likely to be highly cited in the first three to five years, but after that, more novel work fares better.[29] Second, and perhaps more relevantly for returns to innovators, more novel work is published in journals that have lower Journal Impact Factor.[30] In a similar vein, Calcagno et al, 2012, study publications in over 900 journals in the biological sciences between 2006 and 2008. They show that the citations of initially rejected papers were higher than those that were accepted at their first journal.[31]

Of course, the modeling results above do not imply that agents should avoid novelty. Instead, there is a tradeoff is between creativity and how well it is likely to be executed: projects close to known landmarks may be well done, but of little intrinsic value, while those far from known landmarks may not be so well done. This offers the possibility of a "sweet spot" in the middle. Evidence accords with this. First, Mukherjee et al., 2015, and Uzzi et al., 2013, map citations of scientific publications using metrics for novelty, and show that

[28] In their experiment, they show subjects a series of images. For some subjects, images are repeated while for others, the images are all different. They showed that images that were all different induced negative affect in their subjects (it put them in a bad mood) as it increased uncertainty. Mueller et al, 2011, extend this line of work by showing that much of this bias is unconscious: people perceive themselves as open to new ideas to a greater extent than is actually true.

[29] In the first two years after publication, highly novel papers are 30% less likely to be in the top 1% of cited papers, whereas fourteen years after, they are 40% more likely.

[30] For a good related read, see Campanario, 2009, on Nobel prize winning contributions that were rejected by academic journals.

[31] This could alternatively be due to authors responding to rejection by improving their papers.

academic impact is maximized by combining some novelty with already known components. Those who eschew the known for more innovative ingredients at that margin, or who do not have enough novelty, tend to suffer.[32] As a direct example of a non-monotonicity between familiarity and returns, Boudreau et al, 2016, study referee evaluations of endocrinology grant proposals. (The referees are randomly assigned.) They find an inverted U-curve in how ratings relate to the intellectual distance between the evaluator and that embodied in the proposal. Those close to the rater's knowledge, and those far away were rated poorly compared to those of intermediate distance.[33]

The second part of the argument is that incentives are more successful when agents have a less creative focus. There is a small experimental literature in economics on the impact of offering incentives in creative settings, where the creativity of the setting varies.[34] These studies offer subjects the opportunity to take a more or less creative option, and provide some subjects with financial incentives based on a measure of success. The results of the experiments suggest that pay for performance only works when agents do less creative tasks (Charness and Grieco, 2014, Erat and Gneezy, 2016). The closest experimental test is Kachelmeier et al., 2008, where subjects are offered monetary rewards for completing rebus puzzles. Incentives are offered both for more puzzles and for more creative solutions to those puzzles. They note that the highest weighted productivity arose when subjects were offered incentives to produce quantity over more creative options. This reflects the tradeoff here where narrowing assignments in a less ambitions way ultimately improves performance.[35]

For a real world setting outside academia, Balsmeier et al, 2016, study the impact of independent corporate boards on innovation. It is well known that independent boards often implement practices that reduce agency problems, among them measures to make

[32]Their preferred interpretation is that known ingredients are necessary in order to implement the research. Henderson and Clark, 1990, make a similar point.

[33]Note also from the Siler et al., 2015, data in Figure 1 that reviewers were successful in rejecting many papers that subsequently has very few cites: these may have been those that offered little advance over existing knowledge.

[34]Our interest here is in studies where incentives are offered, but where creativity varies. There is a large and inconclusive literature on whether incentives in creative settings work (Mehta et al, 2017, Erat and Gneezy, 2015) We also do not address the literature on using tournaments or competitions to induce creativity. See Bradler et al, 2016, as a recent example.

[35]A useful caveat here is the limited time frame over which effects are measured. Many creative activities percolate over months or years, and it is not clear whether the immediate influences of the experimental studies arise over these longer time frames. Related to this, note that a study by Kachelmeier et al., 2015, found that resampling agents a week after an initial intervention reversed their results: those agents who were placed on incentives performed worse during the experiment, but were more creative a week later.

executives more financially accountable for their performance. They study the impact of board independence on the number and kind of patents that are filed by companies. Two results arise. First, the number of patents rises. Second, it does so only in "more crowded and familiar areas of technology" (p.536). This possibility of better, but less ambitious, outcomes generated by incentive provision is the central message of the paper.

Creativity is often discussed in the arts. Here there is, not surprisingly, little formal evidence. It is clear, however, that there are concerns in some quarters about the limited creative reach of its participants. For example, see Saltz, 2014, on contemporary painting, in a piece whose title, "Why Does So Much New Abstraction Look the Same?", makes the theme clear. Or Roelstrade (2012): "As the art world has got bigger, art itself has somehow become smaller... An important factor in this crippling development has been ... a growing reliance on ever more specialized types of cultural literacy on the part of the beholder - the kind of literacy or meta-linguistic fluency that allows one to really only make oneself fully understood by an alarmingly shrinking audience, in a depressingly small corner of the world (that is to say, the art world)".

Related Theoretical Literature The closest theoretical work is on radical and incremental innovations in research and development settings. In that literature, firms seek to locate their - lets say - product choice on the part of the landscape that has the greatest height. To identify this, they choose between a marginal innovation and a more radical option. By assumption, the marginal change is free to pursue, while radical change is exogenously cost. In these settings, the marginal option can be preferred, sometimes in an inefficient way. In Jovanovic and Rob, 1989, firms choose between implementing an existing project or further developing it. Other firms can potentially mimic, in an efficient way. However, by assumption, ideas can only mimicked if they are being further developed. Because firms cannot charge other firms for this, they implement their ideas too soon, and research moves more incrementally than is efficient. Jovanovic and Rob, 1990, address the distinction between radical and incremental change in a more parametric way. They do not offer a model endogenizing the costs and benefits of incremental and radical change, but instead characterize period of cycles between the two when the benefits of radical change and the costs lie in an intermediate range. Finally, the paper that is formally closest is Gargfagnini and Strulovici, 2015, where agents choose between local and more radical innovation. Once again, the cost of radical change is exogenously higher, and this cost is traded off against the option value of discovering better than current practice. As here, they also model innovation opportunities through a landscape governed by Brownian motion.

This paper varies from the previous work in many ways. First, our focus is on the agency problem that may generate a desire for more local innovation, in ways that are consistent with the literature. Second, here only novelty is rewarded - there is no return to "been there, done that". By contrast, each of the papers above allows new innovators to free ride off previous innovations, by replicating the best innovation currently attained. Here we are interested in settings where there is little payoff to doing this, yet it happens. This is implemented here by the combination of assuming that firms own their ideas and assuming payoffs are based on reducing uncertainty rather than trying to find this highest point on the landscape.

6 Another Interpretation

The last subsection showed how an inability to identify success caused the firm to place more weight on the alternative instrument, inputs, such that restricting creativity becomes more likely. Here we address the possibility that the likelihood of a contribution being recognized depends on its novelty. We do this for two reasons. First, to allow another interpretation that does not involve the dichotomy between inputs and success. Second, to show an outcome where fields can become temporarily fallow.

To do this, consider a case where there are no inputs (so $\Gamma = 0$). Assume that if the agent is successful, as above she receives an incremental reward of $[S(r) - \frac{\overline{w}}{\pi(e^*)}]$, but only receives it with probability $z(r)$, where now $z'(r) < 0$. While the agent is rewarded for success with probability $z(r)$, we assume that the payoff to the principal is $\lambda z(r)S(r)$, where $\lambda > 1$. The interpretation of $\lambda > 1$ is that agents are more short lived than are firms, and success is only recognized later. We formalize this below. When choosing effort, the agent does so to maximize $\pi(e)z(r)[S(r) - \frac{\overline{w}}{\pi(e^*)}] - e$, so that $\pi'(e^*)z(r)[S(r) - \frac{\overline{w}}{\pi(e^*)}] = 1$. Here $\frac{d\hat{U}(e,r)}{dr} = \lambda\pi(e^*)S'(r) + \hat{U}_e\frac{de^*}{dr}$, and

$$\frac{de^*}{dr} = \frac{\frac{z'(r)}{z(r)}[S(r) - \frac{\overline{w}}{\pi(e^*)}] + S'(r)}{H_1}, \tag{20}$$

where $H_1 = -\frac{[S(r) - \frac{\overline{w}}{\pi(e^*)}]}{\pi'(e^*)} + \frac{\overline{w}}{\pi^2(e^*)} < 0$ if the the second order condition holds. Once again, this cannot a priori be signed, but not because of the distinction between inputs and success. Instead, because agents are concerned that taking ambitious projects will not be recognized, $z' < 0$. Once again, limited ambition may be used to encourage effort.

Fallow Periods The purpose of the interpretation above is not simply to show that other assumptions are consistent with a role for limited creative ambition. Instead, it allows the possibility of fields becoming fallow, where agents choose not to search for some period of time, but after some time has elapsed, they will begin searching again. To see this, consider a particular interpretation of λ above, generated by the assumption that in any period where the agent has been previously successful but unrecognized, there is a probability $z(r)$ that recognition occurs. (Once it is recognized, it remains recognized.)

There are two issues that arise with delayed recognition. First, as above, the agent may be concerned that she does not get paid. Second, a firm may hire an agent to work on a project that is later recognized to have been already solved. To address this second influence, assume that the return to success always goes to the earliest innovation, where any later innovator who explored the same partition gets nothing. (In reality, returns are lower if someone else was shown to have got there first.)

The implication of this is that we now need to distinguish between exploring a terrain that no one has attempted to map, and one where previous attempts have been made. First, for previously unexplored terrain, when choosing effort the expected return to the agent is as above, while the firm's expected return is

$$S(r)\left(z(r)+\delta(1-z(r))z(r)+[\delta(1-z(r))]^2 z(r)+....\right)=\frac{S(r)z(r)}{1-\delta(1-z(r))}, \tag{21}$$

and so, with the notation above, $\lambda=\frac{1}{1-\delta(1-z(r))}$. (Note that the firm need not worry about a later agent tackling the same problem as the earliest firm gets the returns.)

The case where someone tried before is more complicated. Here the fear is that it might retrospectively be recognized to be a success. This can give rise to more complicated dynamics as the efficient benchmark is no longer to always take the largest remaining partition. The purpose here is not to describe that more complicated equilibrium but instead to show how this can allow the possibility of a field going temporarily fallow.

To see this, consider a simple setting where only projects up to size $\frac{1}{4}$ are ever attempted. The agent will do so if no one has previously attempted the project.[36] However, if a previous agent has already attempted it with no sign of success, it could be because the previous agent failed, or because she succeeded and it has not yet been recognized. Let it be τ periods since the previous agent tried. Then the probability that she succeeded (without being recognized)

[36] This outcome is defined by $\lambda\pi(e^*(\frac{1}{8}))S(\frac{1}{8})-e^*(\frac{1}{8})<\overline{U}<\lambda\pi(e^*(\frac{1}{4}))S(\frac{1}{4})-e^*(\frac{1}{4})$, where as above $e^*(.)$ is the equilibrium choice of effort for a partition size.

is given by

$$p(\tau; r) = \frac{\pi(e^*)(1-z(r))^\tau}{\pi(e^*)(1-z(r))^\tau + (1-\pi(e^*))}. \tag{22}$$

Note that $p'(\tau) < 0$: the more time that has elapsed, the less likely was success. Furthermore, $p(\tau) \to 0$ as $\tau \to \infty$.

Any firm attempting to map that point is only rewarded if the previous agent was not successful, and so the current agent's expected pay is decreasing in $p(\tau)$. As a result, an agent may be unwilling to explore directly after another agent has tried a project, as $p(1)$ is high.[37] Let τ^* be the smallest time period such that

$$\overline{U} \leq (1 - p(\tau^*))\lambda\pi(e^*(\tau^*))S(\frac{1}{4}) - e^*(\tau^*). \tag{23}$$

Then the model has the feature that after every (failed) attempt at the project, that partition lies fallow for τ^* periods before someone tries it again. Note also that $p(\tau; r)$ is increasing in r, so that this possibility of a field going fallow is more severe for the most creative activities. As such, the result suggests that local innovation does not go fallow, but projects which are most creatively ambitious are those which to temporarily dead.

7 Extensions

Here we consider the role of a variety of assumptions used above.

When Creative Projects are Less Likely to Succeed The model above assumes that the probability of success is independent of creative ambition. This allows us to cleanly distinguish between the efficient benchmark of maximal creativity and the more limited creativity outlined above. In many settings, more creative projects are harder to pull off. Here we consider the robustness of the results to this extension.

Let the probability of success be now given by $\pi(e, r)$, where $\pi_e(e, r) > 0$ and $\pi_r(e, r) < 0$. With this extension, $\frac{d\hat{U}(e^*,r)}{dr} = \pi(e^*)S'(r) + \pi_r S(r) + \hat{U}_e \frac{de^*}{dr}$, and $\frac{de^*}{dr} = \frac{(\pi_e+g)S'(r)+(\pi_{er}+g_r)\left(S(r)-\frac{\overline{w}}{\pi}\right)}{H}$. Let r^{***} be efficient creativity in the absence of the agency problem.[38] Then a redefined analog to Assumption 2 is that effort is below its efficient level at r^{***}. This extension then has similar qualitative features to above as $\frac{d\hat{U}(e^*,r)}{dr}$ has the same influences as above: all it does is evaluate it at a partition potentially different from the largest one. Once again, restricting creativity may be efficient.

[37] This will arise if $\overline{U} > (1 - p(1))\lambda\pi(e^*(1))S(\frac{1}{4}) - e^*(1)$.

[38] This is defined by $\pi_r(e^{**}, r^{***})S(r^{***}) + \pi(e^{**}, r^{***})S'(r^{***}) = 0$, where e^{**} is defined by $\pi_e(e^{**}, r^{***})S(r^{***}) = 1$.

The dynamics in Proposition 2 also continue to hold, but with one necessary condition. Define $\kappa(k)$ as above, where κ is the largest project $\frac{1}{2^{\kappa}}$ smaller than $\frac{1}{2^k}$ that is dominated by $\frac{1}{2^k}$. For Proposition 2 to carry over, we additionally need that the most efficient project in the absence of agency concerns in that range remains the largest one, $\frac{1}{2^k}$. If this is so, then Proposition 2 continues to hold, with intervals of limited creativity. Furthermore, it remains the case that fields can get permanently stuck for the reason outlined above. As such, the qualitative results above do not depend on success being technologically independent of creativity.[39]

The Relationship between Inputs and Success In the model above, inputs only play the role of providing signals on effort. Realistically, inputs cause outputs. Here we discuss the relationship between the inputs and success. Assume that there are n inputs, γ_i, $i = 1, 2, ..n$. These are random variables and are independently distributed $\gamma_i \sim N(\gamma, \sigma_\gamma^2)$. Exerting effort makes the choices more precise: specifically we assume that $e = (\sigma_\gamma^2)^{-1}$.

In keeping with the idea that creative success is rare, the mapping between inputs and success is that the agent must perform sufficiently well at each of the n inputs to achieve success. Success is achieved if γ_i is within z of its correct value for all i. For any single Normally distributed γ_i, the probability that the identified choice lies within $[-z, z]$ of the true value is the error function, whose pdf is given by $\Pi(e) = \int_{-\frac{ez}{\sqrt{2}}}^{\frac{ez}{\sqrt{2}}} exp^{-i^2} di$. As the signals are conditionally independent, the probability of success is $\pi(e) = \Pi(e)^n$. The optimal choice of effort for a project of size r is then given by the solution to $max_e E[\int_{-\frac{ez}{\sqrt{2}}}^{\frac{ez}{\sqrt{2}}} exp^{-i^2} di]^n S(r) - e$. Assuming that the second order condition holds, the optimal choice of effort is given by

$$E\left[nz\sqrt{2}exp^{-(\frac{ez}{\sqrt{2}})^2}\Pi^{n-1}\right]S(r) = 1. \tag{24}$$

This has the same features as the reduced form π function above, where the largest residual partition is tackled first, and effort is increasing in the size of the partition.

[39]This extension allows the possibility of one novel outcome: where the optimal contract can involve more creativity than the myopic first best. Adding the agency problem above implies that we need to evaluate surplus at e^*, where $\left(\pi_e(e^*, r^{***}) + \frac{\Gamma_e(e^*, r^{***})}{\Gamma(e^*, r^{***})}\right)\left(S(r^{**}) - \frac{\overline{w}}{\pi(e^*)}\right) = 1$. Then $\frac{d\hat{U}(e^*, r^{***})}{dr} = \hat{U}_e \frac{de^*}{dr}$, where $\frac{de^*}{dr}$ is defined by (12). As above, $\frac{de^*}{dr}$ and $\hat{U}_e$ can be positive. If so, the firm chooses a level of creativity above the myopic efficient level r^{***} to induce more effort. While this is novel, its intuition is constructive for addressing its empirical importance. This outcome arises if rewarding on success is so valuable relative to rewarding on inputs that the firm wishes to emphasize this. While logically consistent, we have not focused on this here for the reason that much of the empirical motivation for the paper is the difficulty of rewarding success. As a result, we do not emphasize it.

The only remaining issue is to map the information structure into a signal I. Assume that $I = 1$ if $\zeta > 0$ where $\zeta = e + \epsilon_I$, where $\epsilon_I \sim U[-qr, qr]$, where q is large enough that the agent believes that the solution is interior, so $\Gamma < 1$.

Non-contractible project choices So far, we have assumed that the firm can direct the project choice of the agent, through say how it funds scientific proposals. In many creative settings, this could be difficult. As an alternative, consider the case where r cannot be contracted upon, but instead is chosen by the agent. What matters then is what can be contracted upon. If S can be contracted upon, then there is a simple forcing contract where the agent is only paid $w^*_{\Lambda I}$ if and only if success yields a specific payoff, namely the payoff from the project of optimal creativity. At the opposite extreme, consider the benchmark where all that can be contracted upon is "success", $\Lambda \in \{0, 1\}$, where a wage $w_{\Lambda I}$ is paid if any success is achieved.

The agent now chooses r to maximize $\pi(e)S(\hat{r}) - e$, where $\hat{r}$ is now a belief held by the firm about the choice of r made by the agent. The agent's choice of r is given by

$$\frac{d\hat{U}(e^*, r)}{dr} = \pi(e^*)\Gamma_r(e^*, r)(w_{\Lambda I}(\hat{r}) - \overline{w}) + \hat{U}_e \frac{de^*}{dr} = 0, \tag{25}$$

where $w_{\Lambda I}(\hat{r})$ is the optimal bonus based on a belief $\hat{r}$. Here the likelihood of attaining the bonus is affected by r via Γ_r instead of $S(r)$. We have not signed this above. However, the leading cases alluded to above are ones where $\Gamma_r = 0$. If this is the case, and $w_{\Lambda I} > w_0$, then

$$\frac{de^*}{dr} = \frac{g_r(w_{\Lambda I}(\hat{r}) - \overline{w})}{H} < 0, \tag{26}$$

and the agent chooses the least creative option possible. In equilibrium, the wage is then based on surplus for that smallest partition. Yet remember that this is the outcome conditional on $w_{\Lambda I}(\hat{r}) > \overline{w}$, and so incentives are provided. There is one alternative: to simply offer no incentives, and pay independent of success. This involves zero effort but at least generates no reason to restrict creativity, with payoff $\pi(0)S(r^{**})$.

This is a stark outcome - either give no incentives, or take the least creative option - caused by a stark assumption, namely, that the contract cannot distinguish between success on minimally important projects and success on the efficient project. The model can naturally be extended to a setting where the true payoff is probabilistically identified, and the willingness of the agent to choose an ambitious project is increasing in the probability of the true payoff being realized. However, once again this section involves the agent potentially being less creative than optimal, not for the reason that it induces more effort, but as a distortionary response to the fact that ex post success cannot be contracted upon. We have ignored non-contractibility of creative ambition to avoid this additional distortion that is not the focus of the paper.

Noisy Inferences Above we simply assumed that $g_r < 0$. There are a number of natural (and well known) agency settings where the spirit of this outcome arises:

- In the traditional career concerns model of Holmstrom, 1999, agents are rewarded on beliefs of their perceived ability. This perception is based on a Bayesian estimate that depends on measures of performance. As an example, assume that performance on the inputs is $y = f(e) + a + \zeta(r)$, where a is the ability of the agent, $f' > 0, f'' < 0$, and $\zeta(r) \sim N(0, b + cr)$, for $b \geq 0, c > 0$. Hence inferences are noisier for more creative projects. Assume that ability is symmetrically unknown: $a \sim N(\overline{a}, \sigma_a^2)$.

  In the model above, the agent was rewarded on the outcome of the two binary signals in a complementary way. Assume in the spirit of the setting above that the agent is now compensated according to the product of success and expected ability: $w(\Lambda, y) = \beta_0 + \beta \Lambda E(a|y)$, with $\beta > 0$. Then the statistic on which the agent is rewarded in equilibrium is $\Gamma(e, r) = E(a|y) = \overline{a} + \frac{\sigma_a^2(y-(\overline{a}+f(\hat{e})))}{\sigma_a^2+b+cr}$, where $\hat{e}$ is expected effort. With these assumptions, $\Gamma_e = \frac{\sigma_a^2 f'}{\sigma_a^2+b+cr} > 0$, $\Gamma_{ee} = \frac{\sigma_a^2 f''}{\sigma_a^2+b+cr} < 0$, $\Gamma_r = 0$, and $\Gamma_{er} = g_r = -\frac{\sigma_a^2 f' c}{(\sigma_a^2+b+cr)^2} < 0$, as required.

- Many agents are rewarded discretely when their performance passes a threshold. Tenure for an academic would be the most obvious example, and many research awards are based on contests. The canonical representation of this is the tournament model of Lazear and Rosen, 1980. In their baseline model, two workers compete for a promotion that carries a prize. Incentives are generated by the prospect of marginally exceeding the standard needed in expectation to win. Adding noise to the evaluations of agents results in lower incentives, as it makes it less likely that extra effort exerted becomes marginal. (Such a relationship between noise and performance has been empirically verified in Delfgauuw et al.'s, 2015, study of tournaments in a retail sector setting.) Then if more creative actions involve more noisy evaluation, incentives fall.

Myopia We have assumed that agents live for only one period. This simplifies because it implies that the agent chooses the partition that is myopically efficient each period, taking account of her incentives. Myopia is necessary for the characterization of the dynamics above. It is, however, not necessary for creativity being limited for agency reasons, and indeed in some cases forward looking agents can exhibit even less creativity than their myopic counterparts. We show such an example in the Appendix.

Known End Points We have assumed above that there are end points with known height when the game begins. One alternative assumption is where the terrain is a circle where no points are known. The first agent would then choose any point in the circle to identify so that h_0 and h_1 are the same point and 1 is the perimeter of the circle. This offers identical outcomes to above.

A more novel extension is where there is no known end point, and that, in at least one direction, there is a terrain that stretches out infinitely. Consider the case where the terrain is infinite in one direction, from 0 to ∞. This is the setting studied by Garfagnini and Strulovici, 2015. Let $R < \infty$ be the agent's preferred choice of partition when faced with an infinitely sized terrain. Then the first agent chooses point R. This determines a partition from 0 to R, and the results of the previous section hold, but with one additional option for the next agent: to choose $r = R$ again, by locating at point $2R$. If the agent does this, there is a stationary outcome where every agent chooses R *ad infinitum*. This is reminiscent of the outcome in Jovanovic and Roy, 1990, where the returns to innovation are sufficiently high. If R is not the preferred choice of the agent moving after the first R partition is solved, then the results of the model above hold in spirit, where there is an interval of agents taking less ambitious projects than R, but when they get sufficiently small, another $r = R$ is chosen and the next cycle begins.[40]

Varying the Return to Success A natural question to ask here is how the tendency towards limited ambition is affected by the returns to success. So, for example, are fields with big enough returns proof from these problems? If so, these issues may be more prevalent in mature fields where larger projects are already completed. To address this, consider an exercise where instead of the terrain being over $[0, 1]$, it is over $[0, R]$ and R is varied, but where the relative size of partitions remains constant. In the continuous r case of (10), the effect of marginally increasing partition sizes is given by

$$\frac{d^2\hat{U}}{dr^2} = \pi(e^*)S''(r) + \pi'(e^*)S'(r)\frac{de^*}{dr} + \frac{d\hat{U}_e}{dr}\frac{de^*}{dr} + \hat{U}_e\frac{d^2e^*}{dr^2}. \tag{27}$$

[40] The final case is where the terrain is infinite in both directions. Then the first agent randomly chooses a point, and after that the results for the case above hold, except that agents who choose the R can choose either left or right.

This cannot be unambiguously signed. Without even considering the final two terms, it should be clear why. Consider increasing the size of all problems proportionately. Then the value of the bigger partitions in absolute terms have increased. All else equal, this makes the larger ones more desirable: this is the first term in (27). However, failing to succeed is more costly when when returns go up: this is reflected in the second term in (27). If $\frac{de^*}{dr} < 0$, this makes limited ambition even more desirable. Without more information on the elasticity of effort compared to the returns to more success, the relative impact of these cannot be signed. As a result, we cannot make unambiguous conclusions about the impact of more valuable fields on limited ambition.

Midpoints The outcome that allows us to characterize dynamics above is that agents choose the midpoint of their partitions. We showed above that the surplus of a discovery is maximized at the midpoint, but simply assumed that the agent's performance on the inputs only depends on r rather than its location in that space. One could imagine, however, situations where an agent's performance on the inputs might be most informative if located very close to one of the end points of the partition.

To see an alternative, consider a case where the agent need not do so. Then the first agent is faced with a single partition $[0, 1]$ and chooses not to locate at $\frac{1}{2}$ but (without loss of generality) at a point $\rho_1 \leq \frac{1}{2}$. Agents continue at that point until one succeeds. The next agent is then faced with two partitions, $[0, \rho_1]$ and $[\rho_1, 1]$, and locates at point ρ_2. The most ambitious project is $[\rho_1, 1]$, but the analysis above can easily be carried over such the agent may prefer to take a smaller partition. After success at that point, the next agent has choices of either $[0, \rho_2]$, $[\rho_2, \rho_1]$ and $[\rho_1, 1]$, or $[0, \rho_1]$, $[\rho_1, \rho_2]$ and $[\rho_2, 1]$, and so on. This offers the same insights on the possibility of limited ambition, but its dynamics are complicated in a way that does not allow the explicit characterization of Proposition 2.

8 Conclusion

Much is said about creativity, but often with little formal structure. The purpose of the paper has been both to interpret the evidence on creativity through the lens of agency theory, and more positively, to offer the idea that creative ambition may be an important tool in considering the tradeoff between creativity and incentives. Outsiders to many professions often wonder how its members can spend so much time discussing what appear to be minutia. The lens of agency offers an interpretation of this, where the difficulty of evaluating truly innovative work leads fields to spend resources mining further and further into minutia.

References

Amabile, Teresa, "Motivating Creativity in Organizations: Doing What You Love And Loving What You Do", *California Management Review*, 1997, 40(1), 39-58.

Balsmeier, Benjamin, Lee Fleming and Gustavo Manso, "Independent Boards and Innovation", *Journal of Financial Economics*, 2017, 123, 536-557.

Boudreau, Eva Guinan, Kevin J., Karim R. Lakhani and Christoph Riedl, "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance and Resource Allocation in Science", *Management Science* , 2012, 62(10), 2765-2783.

Bradler, C., S, Neckermann, and A.J. Warnke, "Incentivizing Creativity: A Large-Scale Experiment with Tournaments and Gifts", 2016, ZEW Discussion Papers, No. 16-040.

Charness, Gary, and Daniela Grieco, "Creativity and Financial Incentives", 2014, mimeo, UCSB.

Calcagno, V., E. Demoinet, K. Gollner, L. Guidi, D. Ruths, and C. de Mazancourt, "Flows of Research Manuscripts Among Scientific Journals Reveal Hidden Submission Patterns", *Science*, 2012, 6110, 1065-1069.

Campanario, J., "Rejecting and Resisting Nobel Class Discoveries: Accounts by Nobel Laureates", *Scientometrics*, 2009, Vol. 81, No. 2, 549-565.

Deci E. L., and Ryan R. M., *Intrinsic Motivation and Self-determination in Human Behavior*, 1985, New York, NY: Plenum Press.

Delfgaauw, J., Robert Dur, Arjan Non, and Willem Berbeke, "The Effects of Prize Spread and Noise in Elimination Tournaments: A Natural Field Experiment," *Journal of Labor Economics*, 1999, vol. 3, issue 3, pages 521-569.

Eidelman S., Crandall C. S., and Pattershall J., "The Existence Bias," *Journal of Personality and Social Psychology,* 2009, 97, 765-775.

Erat, S., and Gneezy, U., "Incentives for Creativity," *Experimental Economics*, 2016, 19, 269-80.

Fiske, S., and S. Taylor, *Social Cognition*, 1991, New York: McGraw-Hill.

Garfagnini, U. and B. Strulovoci, "Social Experimentation with Interdependent and Expanding Technologies," *The Review of Economic Studies*, 2016, vol. 83, issue 4, pages 1579-1613.

Hall, R.H., "Professionalization and Bureaucratization", *American Sociological Review*, 1968, 70(2), 92-104.

Hennessy, Beth, "The Social Psychology of Creativity: Effects of Evaluation on Intrinsic Motivation and Creativity of Performance", *Multiple Perspectives on the Effects of Evaluation on Performance*, 2001, Springer, Boston, MA.

Holmstrom, Bengt, "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, 1999.

Jovanovic, B., and R. Rob, "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 1989, vol. 56, 569-582.

Jovanovic, B., and R. Rob, "Long Waves and Short Waves: Growth Through Intensive and Extensive Search," *Econometrica*, 1990, vol. 58, issue 6, pages 1391-1409.

Kachelmeier, S. J., B. E. Reichert, and M. G. Williamson, "Measuring and motivating quantity, creativity, or both", *Journal of Accounting Research*, 2008, 46 (May): 341-373.

Kachelmeier, S. J., L. Wang, and M. G. Williamson, "The Role of Incentives in Sustaining High-Creativity Production", 2015, manuscript.

Lazear, Edward, and Sherwin Rosen (1980), "Rank Order Tournaments as Optimal Labor Contracts", *Journal of Political Economy*, 88:5, 841-64.

Manso, Gustavo, "Motivating Innovation", *Journal of Finance*, 2011, 66, 1823-69.

McAfee, Preston, and Richard Fullerton "Auctioning Entry into Tournaments", *Journal of Political Economy*, 1999, 107, 573-605.

Mehta, R., D. Dahl, and R Zhu, "Social-Recognition versus Financial Incentives? Exploring the Effects of Creativity-Contingent External Rewards on Creative Performance", *Journal of Consumer Research*, 2017, 44(3), 536-553.

Monahan, J.L., S. Murphy, and R. Zajonc, "Subliminal Mere Exposure: Specific, General and Diffuse Effects ", *Psychological Science*, 2000, 11, 462-66.

Mueller, Jennifer, Shimul Melwani, and Jack A. Goncalo, " The Bias Against Creativity: Why People Desire But Reject Creative Ideas", *Psychological Science*, 2012, 23(1) 13-17.

Mukherjee, Satyam, Brian Uzzi, Ben Jones, and Michael Stringer, "A New Method for Identifying Recombinations of Existing Knowledge Associated with High-Impact Innovation", *Journal of Product Innovation Management*, 2015, 10, 1-13.

Reeve, J., and G. Nix, "Expressing Intrinsic Motivation Through Acts of Exploration and Facial Displays of Interest," *Motivation and Emotion*, 1997, 21(3), 237-50.

Roelstraete, Dieter, "Echo Chamber: Is Today's Art Too Self-referential?", *Frieze.com*, 2012, June 1.

Saltz, Jerry, "Zombies on the Walls: Why Does So Much New Abstraction Look the Same?", *Vulture.com*, 2014, June 17.

Shapin, Steven, "Hyperprofessionalism and the Crisis of Readership in the History of Science", *Isis*, 96(3), 2005, 238- 243.

Siler, Kyle, Kirby Lee, and Lisa Bero, "Measuring the effectiveness of scientific gatekeeping", *Proceedings of the National Academy of Sciences*, 2015, 112 (2), 360-365.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones, "Atypical Combinations and Scientific Impact", *Science*, 2005, 342 (6157): 468-72.

Wang, J., R. Veugelers, and P. Stephan, "Bias against novelty in science: A cautionary tale for users of bibliometric indicators", *Research Policy*, 2017, 46, 1416-1436.

Taylor, Curtis, "Digging for Golden Carrots: An Analysis of Research Tournaments", *American Economic Review*, 1995, 85, 872-90.

Wilson, James, *Bureaucracy*, Basic Books, New York, 1989.

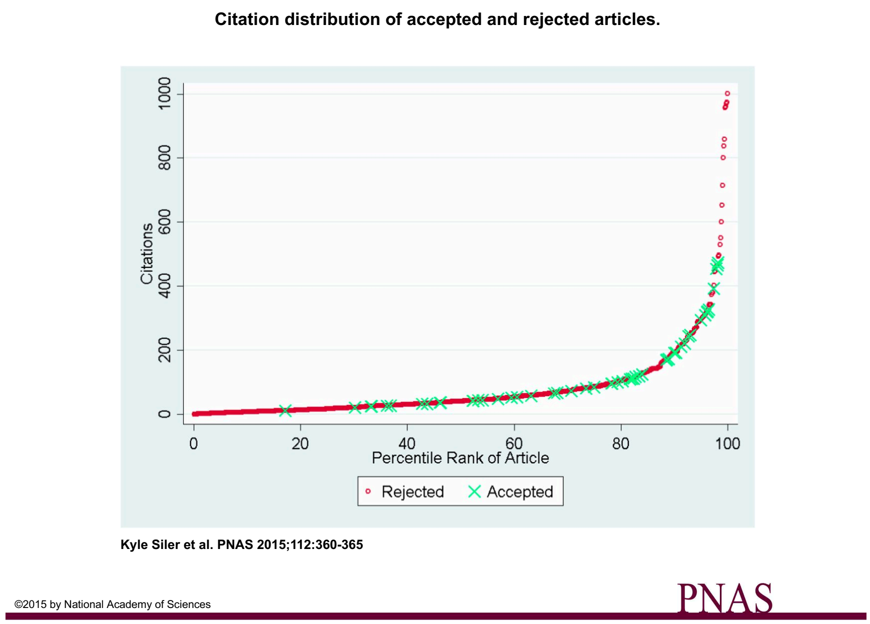

Figure 1: **Citations of Articles Submitted to Medical Journals, taken from Siler et al, 2015.**

Appendix **Derivation of $S(r)$:** For location j, let it be in a partition $[r_1, r_2]$ where r_1 and r_2 are the nearest known neighbors, and let $r = r_2 - r_1$. Because success is either fully revealed, or not revealed at all, any information attained on h_j has value only to the point of known neighbors: beyond that it has no information value. Consider any interval $[0, r]$, and note that for any point i, there are two signals on it, h_0 and h_r, where $h_i \sim N(h_0, i\sigma_\epsilon^2)$ and $h_i \sim N(h_r, (r-i)\sigma_\epsilon^2)$. Standard OLS updating yields $\sigma_i^2 = \frac{i(r-i)\sigma_\epsilon^2}{r}$. Computing this over the range 0 to r gives the firm's utility over that partition as

$$-\int_0^r \frac{i(r-i)}{r}\sigma_\epsilon^2 di = -\frac{\sigma_\epsilon^2 r^2}{6}. \tag{28}$$

Aggregate knowledge entering period t is (28) added over all partitions. Note that (28) is convex in r, so that the optimal choice involves splitting any partition in two equal parts. Doing so implies a value of success for a project of ambition r given by

$$S(r) = \frac{\sigma_\epsilon^2 r^2}{6(1-\delta)} - 2\frac{\sigma_\epsilon^2 (\frac{r}{2})^2}{6(1-\delta)} = \frac{\sigma_\epsilon^2 r^2}{12(1-\delta)}. \tag{29}$$

This is summing the residual variance in both the interval $(r_1, r_1 + \frac{r}{2}]$ and $(r_1 + \frac{r}{2}, r_2)$.

Proof of Proposition 2: Let κ be defined as $min_{(\kappa-k)}$ such that $S(\frac{1}{2^k})[\pi(e^*(\frac{1}{2^k})) - \frac{\pi(e^*(\frac{1}{2^\kappa}))}{(\kappa-k)^2}] \geq e^*(\frac{1}{2^k}) - e^*(\frac{1}{2^\kappa})$. First note that these intervals are of finite length. To see this, consider the case where the agent chooses between a project of size $\frac{1}{2^k}$ and something smaller. To show that there is a project small enough that this is preferred to, consider an upper bound on the agent's utility for a smaller project of size $\underline{k}$. As the agent receives expected surplus, an upper bound is $S(\frac{1}{2^{\underline{k}}})$. Note that $S(\frac{1}{2^{\underline{k}}}) \to 0$ as $\underline{k} \to \infty$ and so there exists a value of $\underline{k}$ which is dominated by k.

The first partition of size $\frac{1}{2^k}$ has been completed. By definition, there must be another partition of that size available. In the completed partition of size $\frac{1}{2^k}$, there are $2^{\kappa-k} - 2$ potential partitions up to size $\frac{1}{2^{\kappa-1}}$. If there is a smaller partition than $\frac{1}{2^k}$ but larger than $\frac{1}{2^\kappa}$ available, the agent will attempt the smaller one. However, by definition a smaller partition is always available, and so $2^{\kappa-k} - 2$ successes are necessary before a project of size $\frac{1}{2^k}$ is attempted.

Now consider the expected number of agents who attempt projects of limited ambition. The probability of success is $\pi(e^*(\frac{1}{2^i}))$ for a project of size $\frac{1}{2^i}$. As a result, the expected number of agents that will attempt the project is $\frac{1}{\pi(e^*(\frac{1}{2^i})}$. Furthermore, there are 2^i partitions of this size, so the expected number of periods before solving all i sized partitions in the interval is $\frac{2^i}{\pi(e^*(\frac{1}{2^i}))}$. But there are $\kappa - k - 1$ partition sizes that have to be succeeded at before the interval ends, and so the expected number of agents (periods) before the interval ends is $\sum_{i=1}^{\kappa-k-1} \frac{2^i}{\pi(e(\frac{1}{2^{k+i}}))}$.

This determines the end of the interval that arises after the completion of the first partition of size $\frac{1}{2^k}$. There are a further 2^{k-1} partitions of size $\frac{1}{2^k}$. As soon as an agent is successful at each of those, the same interval as above arises again in each case. Incentives are identical, and so the same interval arises.

Furthermore, note that as κ be defined by $S(\frac{1}{2^k})[\pi(e^*(\frac{1}{2^k})) - \frac{\pi(e^*(\frac{1}{2^\kappa}))}{(\kappa-k)^2}] \geq e^*(\frac{1}{2^k}) - e^*(\frac{1}{2^\kappa})$, taking project $\frac{1}{2^k}$ offers the lowest utility among all partitions between sizes $\frac{1}{2^k}$ to $\frac{1}{2^{\kappa-1}}$. As effort must be lowest for that partition this also implies that expected compensation is lowest.

Notice that the agent's reservation utility is ignored here. This is because the agent has a choice of projects $\frac{1}{2^k}$ and smaller. If a smaller option is available, this implies that a previous agent succeeded at $\frac{1}{2^k}$ before, so $U(k) \geq \overline{U}$. But if all projects up to κ are preferred to k, then these also offer utility higher than the reservation utility, and so we can ignore it.

At the end of the 2^k intervals initiated by the solution to a partition of size $\frac{1}{2^k}$, all partitions up to size $\frac{1}{2^{\kappa-1}}$ have been completed, and the agent considers whether to attempt a project of size $\frac{1}{2^\kappa}$. If the utility from that is below $\overline{U}$, the game ends. Otherwise, she attempts it and if she succeed, the same game as above occurs except where κ substitutes for k above.

Compensation and Creativity: The following example involves the most creative being worst paid:

- Consider the discrete case above, where the agent exerts effort $e = 1$ only on projects of size ρ_1 to ρ_2. This requires $1 < \rho_3$ but $r^* \in [\rho_1, \rho_2]$. Then if $\frac{1}{2^{\bar{k}}} \leq \rho_2$, the last project attempted in equilibrium has effort exerted.

- Then if $U^*(\frac{1}{2}) < U^*(\frac{1}{2^{\bar{k}}})$, the lowest utility is attained by the agents who take the most creative actions, namely from $\frac{1}{2}$ to ρ_1.

Optimal Contracting with b^{max}: As the agen t is risk neutral above $\overline{w}$, the optimal contract place the most weight possible on the state that provides most incentives to the agent, but if that is not possible, to redirect the residual to the next most informative state. Without constraints on wages, this involves placing all weight on the state where $\Lambda = 1, I = 1$. However, consider the case where $w^*(1,1) > b^{max}$. Then, instead, $w(1,1) = \overline{w} + b^{max}$ is at its maximum level. The firm then has residual expected revenue μ_0, where $\mu_0 = \pi(e)S(r) - \overline{w} - \Gamma\pi b^{max}$. For small μ_0, this residual is assigned either to the state $\Lambda = 1, I = 0$, or $\Lambda = 0, I = 1$. As μ_0 is spread over all cases where this arises, the bonus possible for $\Lambda = 1, I = 0$ is $\frac{\mu_0}{\lambda\pi(1-\Gamma)}$ and for $\Lambda = 0, I = 1$ is $\frac{\mu_0}{\Gamma(1-\pi)}$. Then to maximize incentives, the firm offers the bonus for $\Lambda = 1, I = 0$ if and only if $\frac{\pi'}{\pi(1-\Gamma)} \geq \frac{\Gamma_e}{\Gamma(1-\pi)}$.

If the wage when μ_0 is paid out is less than b^{max}, then this concludes the optimal contract contract. If, however, it exceeds b^{max}, then the only way to offer more incentives is to pay in the other partially successful state, where the firm now has residual bonus money of $\mu_1 = \pi(e)S(r) - \overline{w} - \Gamma b^{max}$ or $\mu_1 = \pi(e)S(r) - \overline{w} - \pi b^{max}$ depending on which signal is the second most valuable. It then spreads that residual μ_1 in the least informative state where the outcome is 1, and will continue to do so until *that* bonus reaches b^{max}. That occurs if $\pi(e^*)S(r) > \overline{w} + (\Gamma + (1-\Gamma)\pi)b^{max}$. In that case, the firm optimally pays a bonus if either signal is positive, as required.

As b^{max} falls below that level, the firm continues to pay b^{max} as a bonus if there if either $\Lambda = 1$ or $I = 0$, but increases w_0 above $\overline{w}$ for budget balance.

Myopia Consider the following extension. Assume now that agents live for two periods instead of one, but are not overlapping. She discounts at a rate δ.The agent captures expected her surplus over the two periods, so that if she chooses a project of size $\frac{1}{2^{k_1}}$ in her first period, and $\frac{1}{2^{k_2}}$ in her second, her utility is $U(k_1) + \delta U(k_2)$. (Note that the set of available projects in period 2 could depend on whether she succeeds in period 1.) The strategic addition to the model above is that the agent may take a myopically dominated project in period 1, because if she is successful, it opens up an even more desirable project in period 2.

This can result in the agent becoming more or less creative than under myopia. To show the possibility of the latter, we provide an example. Consider the case that an agent is born when only two projects have been solved, one of size $\frac{1}{2}$ and one of size $\frac{1}{4}$. That agent in his first period can choose either a project of size $\frac{1}{4}$ or $\frac{1}{8}$. Assume that $U(\frac{1}{4}) > U(\frac{1}{8})$, so a myopic agent always chooses the larger partition. However, let $U(\frac{1}{16}) > U(\frac{1}{4})$. Then the agent may takes the project of size $\frac{1}{8}$, because if she succeeds, she now has the opportunity to take one of size $\frac{1}{16}$. This project could be sufficiently desirable to make it worthwhile. This opportunity is not possible if she takes $\frac{1}{4}$ in the first period. In both cases, if she fails in the first period, she takes the myopically optimal project of size $\frac{1}{4}$ in the second period. Then if

$$U(\frac{1}{4}) + \delta\pi(e^*(\frac{1}{4}))U(\frac{1}{8}) + \delta[1 - \pi(e^*(\frac{1}{4}))]U(\frac{1}{4}) < U(\frac{1}{8}) + \delta\pi(e^*(\frac{1}{8}))U(\frac{1}{16}) + \delta[1 - \pi(e^*(\frac{1}{8}))]U(\frac{1}{4}), \quad (30)$$

the agent chooses the myopically dominated less creative project. In this way, we have shown that allowing agents to be forward looking not only does not eliminate limited creativity, but can make it more likely.