EXPERT ANALYSIS GROUP DISCUSSION PAPER

Scarcity of Ideas and Optimal R&D Policy

By

Suren Basov, Nisvan Erkal, Deborah Minehart§, Suzanne Scotchmer

EAG 24-1 October 2024

Scarcity of Ideas and Optimal R&D Policy

Suren Basov, Nisvan Erkal, Deborah Minehart§, Suzanne Scotchmer

Abstract

We consider a model of innovation that distinguishes between ideas and innovations. While innovation responds to incentives, ideas are a scarce resource that provide an exogenous constraint on the rate of innovation. We investigate how the optimal reward structure is shaped by the scarcity of ideas. Substitute ideas for innovation in the model arrive to random recipients at random times. By foregoing investment in a current idea, society preserves an option to invest in a better idea for the same market niche but with delay. Because successive ideas may occur to different people, there is a conflict between private and social optimality. The social planner does not observe the arrival of ideas and learns over time about the arrival rate of ideas. We represent the social planner's beliefs using a general continuous density and illustrate that the evolution of beliefs can be tracked by the cumulative hazard function. The reward set by the social planner serves a dual purpose: learning about the arrival rate of ideas and trading off lower delays against lower cost ideas. We show that the optimal reward increases as time passes because the social planner comes to view ideas as more scarce.

JEL Classifications: O34, K00, L00
Keywords: Scarcity of ideas: unknown hazard rate: innovations; real options; rewards to R&D

Scarcity of Ideas and Optimal R&D Policy*

Suren Basov[†] Nisvan Erkal[‡] Deborah Minehart[§] Suzanne Scotchmer[¶]

June 5, 2024

Abstract

We consider a model of innovation that distinguishes between ideas and innovations. While innovation responds to incentives, ideas are a scarce resource that provide an exogenous constraint on the rate of innovation. We investigate how the optimal reward structure is shaped by the scarcity of ideas. Substitute ideas for innovation in the model arrive to random recipients at random times. By foregoing investment in a current idea, society preserves an option to invest in a better idea for the same market niche but with delay. Because successive ideas may occur to different people, there is a conflict between private and social optimality. The social planner does not observe the arrival of ideas and learns over time about the arrival rate of ideas. We represent the social planner's beliefs using a general continuous density and illustrate that the evolution of beliefs can be tracked by the cumulative hazard function. The reward set by the social planner serves a dual purpose: learning about the arrival rate of ideas and trading off lower delays against lower cost ideas. We show that the optimal reward increases as time passes because the social planner comes to view ideas as more scarce.

JEL Classifications: O34, K00, L00
Keywords: Scarcity of ideas; unknown hazard rate; innovation; real options; rewards to R&D

*This paper is dedicated to Suzanne Scotchmer. We thank Dan Cao, John Rust, seminar participants at Georgetown University, University of Exeter, University of Queensland, and conference participants at the 22nd Annual International Industrial Organization Conference for helpful comments. This paper supersedes the first part of two working papers by the second and fourth authors entitled "Scarcity of Ideas and Options to Invest in R&D" and "Scarcity of Ideas and R&D Options: Use it, Lose it or Bank it."

[†]Department of Economics, University of Melbourne; surenbasov@gmail.com.

[‡]Department of Economics, University of Melbourne; n.erkal@unimelb.edu.au.

[§]United States Department of Justice; deborah.minehart@usdoj.gov. The views expressed are not purported to reflect those of the Department of Justice.

[¶]Deceased.

1 Introduction

Economic progress depends on the speed with which innovations emerge. What determines the speed with which innovations emerge? This depends both on the challenges at hand and the incentive structure at play. A predicate for innovating is that the innovator must first have an idea for an opportunity to invest in. A measure of innovation difficulty is the scarcity of ideas for a solution to a problem. However, the innovator must also have the incentive to innovate and this depends on social policy such as patent law and public subsidies for R&D. We investigate the interplay between these two factors and ask how optimal rewards in R&D are related to the scarcity of ideas.

We conceive of an idea as an investment opportunity, and thus distinguish between ideas and innovations. Our purpose is to capture the distinction between an exogenous process that generates ideas for innovations and an incentive structure that influences the decisions to invest in the ideas. We assume that the exogenous process cannot be influenced by incentives as it is determined by the nature of the problems society faces and the creative capacity of the human mind. Policy makers can influence which ideas are turned into innovations through the incentive structures they set. Hence, from a policy perspective, an important economic parameter is the frequency with which ideas for investment occur. Although it is rarely the case that an idea is unique in the sense that nobody thinks of a similar idea, similar ideas may occur at different rates. We define ideas to be 'similar' if they fill a given market niche.[1] We define ideas to be 'scarce' if ideas to fill the given market niche occur at long intervals.[2]

Ideas in our model arrive to random agents at random times. Although ideas are similar in the sense that they all fill the same market niche, we assume that they differ based on the investment cost. Each idea has a random investment cost associated with it, and to create an innovation, the recipient of an idea has to invest the cost.

[1] For example, following the establishment of a prize by the British Parliament in 1714 for the measurement of longitude at sea, there were many alternative methods proposed. The proposed methods would be tested by sailing through the ocean from Britain to a port in West Indies. Eventually, the prize produced two distinct workable solutions.

[2] The terms market and market niche as used in this paper do not refer to antitrust markets.

Investing provides the recipient of an idea with profit, depending on the reward system, but not investing preserves a social option: Another agent with an even better (i.e., lower cost) idea might invest instead. The option that is preserved by not investing is a social option, not internalized by any potential investor. This change in perspective expands the purpose of the R&D reward structure. A good reward structure must not only motivate effort and cover the cost of innovation, but it must also encourage investment in the right idea at the right time. It must ensure that ideas are discarded when it is better from society's point of view to preserve the social option.[3]

Our objective is to show how rewards should reflect the scarcity of ideas, understood as their arrival rate. Progress can be slow in our model, not only because R&D is costly and resources are scarce, but also because ideas for investment are themselves scarce, not only from an individual's point of view, but also for society as a whole. Whether an idea should be used or discarded depends both on the costliness of investing in it and on the rate at which new ideas will arrive. If all ideas were available at the same time, the goal of the social planner would be to find the minimum cost idea. However, that is not possible because ideas arrive at random times. If other ideas are expected to arrive rapidly, it might be best to discard a given idea and preserve the option for society to invest in a better (i.e., lower cost) idea. The social planner can weed out high-cost ideas by offering limited rewards, but she still faces a trade-off between cost and delay. To ensure that the market niche is filled at low cost, she may have to endure a costly delay. However, waiting may be too costly if ideas are not expected to arrive rapidly.

We assume that the arrival of ideas (investment opportunities) as well as their costs are private information. Hence, the social planner does not observe the arrival of ideas prior to the prize being claimed. However, the distribution of the possible costs is common knowledge. We assume that the social planner's R&D policy takes the form of a reward structure. For example, the reward function can represent a patent policy or a prize system. An attractive aspect of such a policy is that it is decentralized in that it does not elicit

[3]In a similar vein, Bryan and Lemus (2017) analyze how policy can be structured to solve directional inefficiencies in R&D.

information beyond the fact of investment. Any innovator with an idea may claim the reward after turning their idea into an innovation.

We consider two models. The first model is a benchmark model in which the scarcity (arrival rate) of ideas is known to the social planner. Ideas rain down on a population of innovators according to a continuous time Poisson process. The Poisson arrival rate is the measure of the scarcity of ideas. The social planner chooses a reward policy to maximize social welfare. An innovator who receives an idea can either invest in the idea and claim the prize, or abandon the idea. We solve for the optimal prize using techniques from dynamic programming. Because the environment is stationary, the optimal policy consists of a fixed reward. An idea is invested if and only if its cost is lower than the reward, that is, the reward operates as a cut-off value for whether ideas are financed. From an investor's point of view, R&D profits are positive due to the scarcity of ideas and the fact that the cost is the investor's private information. From society's point of view, the reward is a screening mechanism that balances the trade-off between cost and delay. The cost of the marginal idea generates a social value equal to the value of waiting for a lower cost idea to arrive. We show that the optimal reward should increase with the scarcity of ideas. Society should be willing to tolerate higher cost in environments where ideas are scarce (the arrival rate is low). Society should also be willing to tolerate higher cost in environments where ideas are more costly (i.e., the cost distribution is less favorable in the sense of stochastic dominance).

In the second model, we consider the more realistic case where the scarcity of ideas is not known. Innovation, by nature, is an activity where potential innovators often do not know whether they will succeed. Hence, that the arrival rate of ideas is not known is an important component of the uncertainty of R&D. What potential innovators know is that failures and dead ends are as much a part of the innovative process as successes are.

In this case, the social planner's beliefs about the true rate of arrival evolve over time and as a result, the optimal policy design is no longer stationary. The social planner's evolving beliefs are the state variable for the dynamic optimization problem that determines the optimal reward at each point in time. Our focus is on the interplay between the social planner's policy and what the planner can learn about the scarcity of ideas. Since the arrival of ideas is private information, the social planner does not directly observe the arrival of ideas. However, the belief updating must account for the fact that some ideas may be rejected due to their high cost. The social planner will update based on the amount of time that passes without the prize being claimed, which depends on the policy set by the social planner.[4] Hence, when the arrival rate of ideas is not known, the policy set by the social planner serves two purposes: It determines which ideas will be invested in as well as how quickly the social planner will learn.

We model the social planner's beliefs using a general continuous density that evolves over time. We show that the evolution of the social planner's beliefs about the arrival rate can be tracked by the cumulative hazard function. The cumulative hazard function serves as a one-dimensional Markov state variable for the dynamic optimization and reduces the complex history of the game to a sufficient statistic that can be used for analysis. We use this insight to derive results about how the beliefs evolve. We show that as time passes with no innovation, the planner becomes more pessimistic about the arrival rate. This can be seen in the fact that the beliefs at a later time are dominated by beliefs at an earlier time in the sense of first order stochastic dominance. We also present differential equations for how the moments of the beliefs (the expected value as well as higher moments) evolve over time. Our results reveal how the reward policy determines the pace of updating. A higher reward policy increases the likelihood of innovation, but it also causes the social planner's beliefs to grow more pessimistic more quickly.

We use our results about the evolution of beliefs to characterize the optimal reward policy in this dynamic environment. Our main result is that optimal rewards should increase with delay in filling the market niche. As the planner becomes more pessimistic about the arrival rate, she should tolerate higher cost in order to reduce further delay. An idea that seemed

[4]That the passage of time can be an indication of the difficulty of a problem or the scarcity of ideas is built into the patent law through the concept of "long-felt need." Long-felt need is one of the secondary considerations for patentability and it is a criterion that can be used to fulfill the non-obviousness requirement of a patent application. Note that making deductions based on how much time has passed without a solution to a recognized problem implies that the policy makers do not know what the arrival rate is.

too costly a year ago will seem more attractive at present because more delay is predicted. In part, this generalizes the result from our benchmark model that when ideas are more scarce, society should tolerate higher costs. However, there is now a dynamic consideration that does not factor into the static model. The social planner understands that setting a higher reward will cause him to learn faster and to grow pessimistic more quickly. This impacts the optimal reward policy, but it does not change the main conclusion that the optimal reward increases over time. As in our benchmark model, we also analyze how a change in the cost distribution impacts the optimal reward policy. We present conditions on the cost distributions such that society should be willing to give a higher prize in environments where ideas are more costly (i.e., low-cost ideas are scarce).

Our paper contributes to the R&D literature in two ways. First, the fact that similar ideas may arrive over time has not been considered in the literature. In distinguishing between ideas and innovations, we follow Green and Scotchmer (1995), O'Donoghue et al. (1998) and Scotchmer (2004). Green and Scotchmer (1995) and O'Donoghue et al. (1998) address environments where ideas are complements, and Scotchmer (2004) addresses environments where ideas serve different market niches. Because ideas are not substitutes in these papers, delay in investment is never optimal. In the model we discuss here, it is because ideas are substitutes that a certain amount of delay should be tolerated. One of the ideas that arrives during the delay may have low cost. Second, our modelling approach implies that the optimal reward structure is increasing over time. We do not know of other papers where the reward policy changes dynamically. The dynamics in our model is driven by the social planner's learning process. Our assumption that the arrival rate of ideas is unknown follows Choi (1991) and Malueg and Tsutsui (1997). However, in these papers, there is no distinction between ideas and innovations. There is an unknown parameter that governs the hazard rate of success in a production function for R&D that is common knowledge among the firms in a race. The hazard rate can take on one of two values and as time passes without success, firms update their beliefs about the hazard rate. In our model, the social planner is learning about the hazard rate at which the population as a whole receives ideas for investment. We assume that the social planner's beliefs are a density with support on R_+.

Our model is a real options model in the spirit of MacDonald and Siegel (1986) and Dixit and Pindyck (1994). An investment is irreversible and could turn out to be a mistake. To avoid mistakes, there is a value to delay. In many real options models, the value of the option is internalized by the firm.[5] In our model, ideas (investment opportunities) accrue to random firms, which means that although waiting is valuable to society, the value of waiting is not internalized by any potential innovator. The problem of the social planner is to ensure that private recipients of ideas preserve socially optimal options.

Our modelling apparatus is reminiscent of search models (e.g., McCall and McCall, 2008) although we do not interpret our random process as search, and there is an important difference between our social planner and a job searcher. In search models, all opportunities arrive to a single searcher who sets an optimal stopping policy. In our model, no individual is likely to receive more than one idea. The social planner who sets the reward policy does not observe who receives ideas and when. Despite this fundamental difference, if the social planner knows the arrival rate of ideas, the optimal reward policy is similar to the stopping rule that emerges in search models. However, our main focus is the more realistic case where the arrival rate of ideas is not known. In this case, to choose an optimal policy, the social planner must infer something about the arrival rate of ideas using the only information available to him, which is the time that has passed without success. Although some search models involve learning,[6] the case where the searcher is learning about the arrival rate of offers has, in general, not been studied in the literature. One exception is Mason and Välimäki (2011), where the seller of an asset faces uncertainty about the demand it faces and learns over time about the arrival rate of buyers to the market. Our environment is different from Mason and Välimäki (2011) because we consider a policy maker who maximizes social welfare instead of private profits, so the level of the reward does not impact social welfare

[5] See, for example, Weeds (2002), where firms may want to delay investment in the hope of gaining a higher payoff in the future.

[6] See, for example, Rothschild (1974) and Rosenfield and Shapiro (1981), where there is learning about the distribution of prices. See also Kohn and Shavell (1974).

directly. Moreover, in Mason and Välimäki (2011), the arrival rate of buyers can take on one of two values, allowing for a one-dimensional state variable. We assume that the social planner's beliefs are a density with support on R_+. We contribute by showing that despite the fact that the support is a continuum, the social planner's problem allows for a one-dimensional state variable. We provide an analytic updating formula and differential equations for all of the moments of the social planner's beliefs.

Our model may be of interest to other researchers studying optimization problems with state variables that are probability distributions. For example, recent papers in macroeconomics by Cao and Luo (2017), Nuno and Moll (2018), and Achdou et al. (2022) address a variety of problems using Kolmogorov forward equations to characterize how a relevant distribution evolves. The technical contribution we make with our real options model where the evolution of the distributional state variable can be characterized with a one-dimensional parameter may have applications in macroeconomics or other areas of economics.

The paper is structured as follows. In Section 2, we set forth our model of scarce ideas. In Section 3, we stipulate the social planner's objective and derive the optimal policy in the case when the rate of arrival is known. In Section 4, we characterize the optimal reward policy when the rate of arrival is unknown and show how the learning process of the social planner will change the optimal policy. We conclude in Section 5 by discussing some ways that the optimal reward policy corresponds to legal institutions. All proofs are relegated to the Appendix.

2 A model of ideas and innovations

We assume there is a market niche that may be filled with an innovation. The social value of solving the problem and filling the market niche is v/r, where r is the discount rate. There is an exogenous process by which potential innovators receive solution ideas for filling the market niche. To innovate, the inventor must first have an idea and then an incentive to invest in it.

We assume that ideas rain down on the population as a whole according to a Poisson process with parameter λ, and we take the parameter λ as a measure of scarcity. If the arrival rate λ is low, ideas are scarce. We interpret λ as representing the fundamental difficulty of the problem, i.e., the difficulty of filling the market niche. Different problems may have different fundamental difficulties, but a particular market niche has a given difficulty. Society may not know what this difficulty is, but it can learn about it over time.

Each idea occurs at a random time, to a random recipient. We assume that there are many potential innovators, and each receives at most one idea. [7] This is an intentionally extreme assumption that highlights the main premise of the paper. Ideas are scarce, not only for society as a whole, but especially from the perspective of any individual. The recipient of an idea can invest in it or discard it. If the recipient invests the idea, the process stops because the market niche has been filled. If the recipient does not invest in the idea, the idea is lost to everyone, including the recipient, and cannot be reclaimed later. For example, an idea may be lost or forgotten because the recipient moves on to other projects.

Each idea has associated to it an R&D cost c that is drawn independently from a fixed distribution. The fact that ideas have different costs means that they are imperfect substitutes for filling the market niche.[8] All else equal, an idea with lower cost will generate higher social benefit $v/r - c$. We assume that the cost distribution is given by a continuous probability density f with $f > 0$ on its support $[\underline{c}, \overline{c}]$ in R_+ and cumulative distribution F. To create an innovation, the recipient of an idea must invest the cost c. We assume that $\underline{c} < \frac{v}{r} < \overline{c}$ so that some, but not all, ideas yield a benefit to society.

The social option preserved by not investing in an idea has value because another idea might be less costly. There is thus a social trade-off between cost and delay. However, this social trade-off is not the private trade-off because the next idea will occur to someone else. Because an innovator does not expect to receive another idea, the innovator might be

[7] Formally, the population of potential innovators is assumed to have N individuals. If an idea arrives, it will hit each innovator with equal probability so that in a time interval Δt, an innovator has probability $\frac{\lambda \Delta t}{N}$ of receiving the idea. As $N \to +\infty$. the probability that an innovator receives two ideas either at one point in time or over time become negligible and can be ignored.

[8] As explained in Agrawal et al. (2021), there are often multiple potential strategies for bringing the same idea to the market. The different costs can be interpreted as corresponding to these different strategies.

too willing to invest in the idea he possesses. Individual innovators are motivated by their private profit, not with the social value of preserving options. The policy challenge is to manage private incentives to invest in a way that is socially optimal.[9]

Like all contracts, R&D incentives must depend on factors that are verifiable. Because ideas are private information, the social planner does not observe the arrival of ideas. A social planner offering a reward does not observe the arrival of ideas that recipients reject, but she does observe whether the market niche is filled. The optimal policy will operate by getting the population of potential innovators to screen their ideas and then to discard those with costs that are too high. We assume that the social planner provides the incentive for investment by offering a reward p to the first innovator to fill the market niche. For example, the reward function can represent a patent policy or a prize system. An attractive aspect of such a policy is that it is decentralized in that it does not elicit information beyond the fact of investment.[10]

An innovator with an idea may claim the reward after turning their idea into an innovation.[11] If the recipient of an idea invests in it, the process stops because the market niche has been filled. Because individual agents must 'use or lose' their ideas, a recipient of an idea with cost c invests if and only if $c \leq p$. We say an idea is 'viable' at time t if it has cost less than the reward $p(t)$. Conditional on an idea arriving, the probability that the idea is viable is $F(p(t))$, where F is the cumulative distribution of the cost density f. The reward policy functions as a time-varying optimal stopping rule such that the investment process survives to time t if there is no viable idea until t. The instantaneous arrival rate

[9]Note that because ideas arrive at random times to random recipients, achieving social efficiency does not imply that society has no ex-post regret. Had the process continued, a better idea with a lower cost could have arrived. Indeed, there are several examples of society not investing in the best technology. Consider, for example, the reciprocating-piston internal combustion engine. The Wankel rotary engine, which was developed after the piston engine, is simpler and arguably superior. However, due to "a pure accident of history," it never challenged the dominance of the piston engine in a real sense (Soctchmer, 2004, p. 57).

[10]One could, in theory, design a mechanism to elicit the information on the arrival of ideas from the innovators. However, in this paper, our focus is on environments where ideas are difficult to verify prior to investment. We take reward structures which are commonly used in the real world and ask how they should respond to the scarcity of ideas.

[11]This is because after investing in the idea and turning it into an innovation, the innovator can either show the innovation to claim the prize or can complete a patent application with the required level of detail.

of a viable idea at time t is $\lambda F(p(t))$. A higher value of λ generates a higher arrival rate because ideas are less scarce. A higher reward $p(t)$ generates a higher arrival rate because ideas with higher costs would be invested. Finally, the cost distribution F affects the arrival rate because conditional on arrival, it determines the probability that the idea which has arrived is viable. For example, when ideas have lower costs in a distributional sense, then the arrival rate is higher because more of the ideas which arrive are likely to be viable.

We consider two models. We treat the first model as a benchmark model where the social planner knows the true value of the arrival rate λ. Society faces uncertainty regarding the arrival time of ideas and their costs, but the underlying environment determined by λ and F is stationary. The social planner solves a dynamic optimization problem to incentivize innovators to invest their ideas based on the cost c. We analyze the model using standard techniques from dynamic programming.

The second model is our main model where the social planner does not know the value of λ. This may be because the social planner has uncertainty about the difficulty of the problem or the expertise of the pool of potential innovators. The parameter is not fundamentally random, but society does not know what it is. The assumption of unknown λ adds another layer of uncertainty (and realism) to the problem. We assume that the social planner's initial beliefs about λ, at the time when the incentive structure starts, are given by a probability density $h_0(\lambda)$ defined on $\lambda \in R_+$. We make the following assumption on h_0:

Assumption 1 *h_0 is continuous and bounded, has finite moments, and $h_0(0) > 0$.*

In the second model, the length of time without arrival of a viable idea is a signal of λ. The social planner knows how long it has been since she first offered an incentive and updates her beliefs to reflect this. We denote the posterior beliefs at time t by $h_t(\lambda)$. Intuitively, a long period with no arrival should make the social planner more pessimistic about λ, causing the posterior distribution on λ to shift toward lower values. However, the posterior distribution of λ must also account for the fact that some ideas are rejected. Thus, the social planner's reward policy is an ingredient to forming a posterior belief at time t. If the social planner offers a higher reward at time t, the chance of innovation is higher, but the social planner will become more pessimistic about the arrival rate going forward if the reward is not claimed. When we characterize the evolution of the beliefs in Section 4, Assumption 1 simplifies the analysis by ensuring that the beliefs $h_t(\lambda)$ converge uniformly over time.[12]

3 Optimal reward policy when the arrival rate is known

As a benchmark, we first consider the case when society knows the true Poisson arrival rate λ. At each instant of time, the social planner chooses a reward policy $p(t)$, knowing that if an innovator receives an idea in that instant with cost $c \leq p(t)$, the innovator claims the reward and the game ends. Otherwise the innovator lets the idea go, and society continues to wait for the niche to be filled.

For a given reward policy $p(t)$, the expected cost of a viable idea based on the cost distribution F is

$$E(c|p(t)) = \frac{\int_{\underline{c}}^{p(t)} cf(c)dc}{F(p(t))}$$

Clearly, the expected cost is higher when the reward policy is higher.

As seen from time 0, the probability that the market niche remains open at time t (the survival probability) is

$$e^{-\lambda\phi(t)} \text{ where } \phi(t) = \int_0^t F(p(s))ds \tag{1}$$

In survival analysis, $\lambda\phi(t)$ is called the cumulative hazard function. It represents the cumulative risk that the market niche will be filled by time t. Given our R&D context, we interpret $\phi(t)$ as a measure of the innovation incentive society has offered up to time t. The greater this is, the more likely innovation is to occur by time t. Notice that $\phi(t)$ is a measure of the effective innovation incentives that have been offered, rather than the raw monetary value. The cumulative innovation incentive always increases over time, and a more generous

[12] Assumption 1 ensures that the social planner's beliefs converge uniformly to a point mass at $\lambda = 0$ in the absence of innovation. Assumption 1 can be relaxed, for example, to allow for a positive lower bound on the arrival rate.

reward policy (up to $\bar{c}$) causes it to grow faster:

$$\frac{d\phi(t)}{dt} = F(p(t)) \tag{2}$$

The social planner's optimization problem can be stated as

$$V = \max_{p(\cdot)} \int_0^\infty e^{-(rt+\lambda\phi(t)))} \left(\frac{v}{r} - E(c|p(t))\right) \lambda F(p(t))dt \tag{3}$$
$$\text{subject to } \frac{d\phi(t)}{dt} = F(p(t))$$

where V is the value of social welfare at the optimal policy. This expression reveals the dynamic trade-off faced by the policy maker. A higher reward policy $p(t)$ at time t would imply less delay since the arrival rate of viable ideas $\lambda F(p(t))$ is higher. However, it would also imply a higher expected cost of innovation, as shown by $E(c|p(t))$, because innovators with higher cost ideas can claim the reward.

Note that the optimal reward policy must satisfy $p(t) \in [\underline{c}, \frac{v}{r}]$. A policy $p > \frac{v}{r}$ would allow ideas with negative social value to be invested resulting in lower social welfare than $p = \frac{v}{r}$. A policy $p < \underline{c}$ generates no investment resulting in lower social welfare than any $p \in (\underline{c}, \frac{v}{r})$. Accordingly, we can assume that $p(t) \in [\underline{c}, \frac{v}{r}]$.

To solve the model, we apply standard techniques from dynamic programming. We approximate continuous time with a discrete time interval Δt, solve the Bellman equation for the discrete time version of the social planner's optimization problem, and then take the limit as $\Delta t \to 0$. The arrival rate $\lambda \Delta t$ is the probability that an idea arrives in the time interval. The arrival rate of a viable idea is $\lambda \Delta t F(p)$. As $\Delta t \to 0$, the discount rate $e^{-r\Delta t}$ is approximated by $(1 - r\Delta t)$.[13]

Using Bellman's principle of optimality, we express the social planner's optimization problem recursively as the sum of a current payoff and a future continuation payoff. The Bellman equation is

$$V = \max_{p \in [\underline{c}, \frac{v}{r}]} \{\lambda \Delta t \int_{\underline{c}}^{p} \left(\frac{v}{r} - c\right) f(c)dc + (1 - \lambda \Delta t F(p))(1 - r\Delta t)V\} \tag{4}$$

[13]Since $lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$, terms involving $o(\Delta t)$ can be ignored in the optimization. As $\Delta t \to 0$, the discount rate is $e^{-r\Delta t} = (1 - r\Delta t) + o(\Delta t)$. The probability that one idea arrives in the time interval is $\lambda \Delta t + o(\Delta t)$ and the probability that two or more ideas arrives is $o(\Delta t)$.

In the current period, the social planner chooses the reward policy p to maximize the sum of the current and future payoffs. The first term on the right-hand side of (4) is the social welfare if a viable idea arrives and is invested. The second term is the social welfare to be realized in the future, if no idea is invested today. The future continuation payoff V is the same as the current continuation payoff V. This reflects the stationarity of the environment. If no viable idea arrives today, the cost and benefit parameters do not change, and so the trade-offs faced by the social planner do not change.

Subtracting V from both sides of equation (4), dividing by Δt, and passing to the limit $\Delta t \to 0$, we obtain the Hamilton-Jacobi-Bellman equation for V:

$$0 = \max_{p \in [\underline{c}, \frac{v}{r}]} \{\lambda \int_{\underline{c}}^{p} \left(\frac{v}{r} - c\right) f(c)dc - (r + \lambda F(p))V\} \tag{5}$$

Note that time t does not appear in (5) as the environment is stationary. Taking the derivative of the right-hand side of equation (5) with respect to p, the first order condition for the optimal reward policy p^* is:

$$0 = \lambda \left(\frac{v}{r} - p^*\right) f(p^*) - \lambda f(p^*)V$$

Since $f(p^*) \neq 0$, this simplifies to

$$V = \frac{v}{r} - p^* \tag{6}$$

Equation (6) shows that the social value of investing the marginal idea with cost p^* is equal to the continuation value of waiting for a better idea to arrive.

Substituting p^* for p in equation (5) and using integration by parts, we have:

$$0 = \lambda F(p^*) \left(\frac{v}{r} - p^*\right) + \lambda \int_{\underline{c}}^{p^*} F(c)dc - (r + \lambda F(p^*))V \tag{7}$$

Using (6) to eliminate V in (7), we can characterize the optimal reward policy in terms of the cost and benefit parameters of the model.

$$\frac{v}{r} - p^* = \frac{\lambda}{r} \int_{\underline{c}}^{p^*} F(c)dc. \tag{8}$$

The left hand side (8) can be interpreted as a social demand schedule of ideas and the right hand side as a social supply schedule. Uniqueness of the optimal policy follows from the observation that the demand schedule is decreasing in p while the supply schedule is increasing in p. The optimal policy is the unique value of p^* that brings the two sides of (8) into equality.

The optimal policy has an intuitive relationship to the the scarcity ideas. As ideas become more scarce (i.e., as λ decreases), the social supply schedule shifts down. The social planner increases the reward policy until (8) holds. That is, the social planner sets a higher reward p^* allowing more costly ideas to be invested. Conversely, when ideas are less scarce, the planner sets a lower reward and so is more selective.

Combining (6) and (8) provides a useful relationship between the value function and the optimal policy:

$$V = \frac{\lambda}{r} \int_{\underline{c}}^{p^*} F(c)dc \tag{9}$$

We summarize these results in Theorem 1 which characterizes the optimal reward policy of the social planner and states how it changes with λ.

Theorem 1 *The optimal reward policy p^* is the unique solution to equation (8). The optimal reward policy is increasing in the scarcity of ideas represented by a decrease in λ. As ideas become more scarce, social welfare decreases.*

Theorem 1 implies that rewards to innovation will be higher in situations where ideas for innovation are scarce. When ideas arrive less frequently, the social planner finds it optimal to set a higher prize in order to reduce the expected amount of delay.

Because the optimal policy is stationary, the cumulative innovation incentive grows linearly over time according to

$$\phi(t) = F(p^*)t$$

Using equation (3), it is straightforward to calculate that

$$V = \frac{\lambda F(p^*)}{r + \lambda F(p^*)} \left(\frac{v}{r} - E(c|p^*) \right) \tag{10}$$

This shows that social welfare takes the form of a delayed investment payoff, $\frac{v}{r} - E(c|p^*)$, where the discount factor for the delay is increasing in the arrival rate $\lambda F(p^*)$ of a viable idea.[14]

In the next section, we develop a model with learning where the optimal policy changes over time. We will show that as time passes without an innovation, the social planner comes to believe that ideas are more scarce. With that in mind, our next corollary records properties of the social planner's optimal policy as the arrival rate λ decreases to 0. These properties follow immediately from (8) and (9).

Corollary 1 *As the arrival rate λ decreases to 0, the social planner's optimal reward policy increases to $p^* = \frac{v}{r}$ and social welfare decreases to $V = 0$.*

The optimality conditions (8) and (10) also reveal how the cost distribution F affects the optimal reward policy p^*. To demonstrate this formally, we consider a second cost distribution G with continuous probability density $g > 0$ on its support $[\underline{c}, \overline{c}]$, and assume that F first order stochastically dominates G. Theorem 2, proved in the appendix, states that when ideas tend to have higher costs in a distributional sense, society should be willing to bear a higher cost p^* to fill the market niche sooner, as waiting for a less expensive idea to arrive is likely to take longer.

Theorem 2 *If G is a cost distribution and if F dominates G in the sense of first order stochastic dominance, then the optimal reward policy is higher under F than under G and social welfare is lower.*

We can see that replacing the cost distribution F with G on the right hand side of (8)) has the same impact on the supply schedule of ideas as an increase in λ. Both shift the supply schedule of ideas up, leave the demand schedule for ideas (the left hand side of (8)) unchanged, and thus decrease p^*. For given λ, the cost distribution determines the scarcity of low-cost ideas. Low-cost ideas are expected to be more scarce under distribution F than under distribution G since F first order stochastically dominates G. Theorem 2 establishes that society finds it optimal to give higher rewards when low-cost ideas are expected to be more scarce.

[14] As is typical of models with a Poisson arrival process, the terms in the denominator, $r+\lambda F(p^*)$, represent an augmented discount rate. See, e.g., Loury (1979) and Lee and Wilde (1980).

4 Optimal reward policy with an unknown arrival rate and privately observable ideas

We now turn to the more realistic case that the arrival rate λ is unknown. Except for the uncertainty about the scarcity of ideas, the model is unchanged. The social planner again chooses a reward policy $p(t)$ at each instant of time. If a viable idea arrives, then the recipient of the idea claims the reward ending the game. Otherwise, the innovator lets the idea go, and society continues to wait for the market niche to be filled. The central challenge introduced by the uncertainty is that the social planner's optimal reward policy will depend on the beliefs about the scarcity of ideas which are changing over time.

The uncertainty about λ changes the trade-off between cost and delay, introducing the possibility of learning. Although the social planner does not observe the arrival of ideas, she does observe whether or not the reward is claimed. Intuitively, if a reward is not claimed, the social planner should infer that the arrival rate is lower than expected. That is, ideas are more scarce.

The reward policy now plays a dual role. A higher reward makes innovation more likely, but it also reveals more information about the scarcity of ideas. If the social planner is conservative and sets $p(t)$ at a very low level so that few ideas are viable, then there is little to learn from the fact that the reward is not claimed. Conversely, if the planner is so generous that every idea is viable, then if the reward is not claimed, the planner can be certain that no idea arrived.

4.1 Evolution of social planner's beliefs

The social planner's beliefs about the arrival rate λ is central to the analysis. The beliefs are the only aspect of the model that changes over time, and as such are a sufficient statistic for the dynamic aspects of the social planner's decision problem. We allow for the prior beliefs $h_0(\lambda)$ to have support on all of R_+ so that the uncertainty about the arrival rate can take very general forms. While we do not model the formation of the initial beliefs, we do analyze how future beliefs $h_t(\lambda)$ at time t reflect the history of the social planner's reward policy.

At time t, the social planner updates her beliefs taking into account the history of the reward policy and the fact that the reward is not claimed. As stated in (1), for any given value of λ, the probability that the market niche remains open at time t is given by $e^{-\lambda\phi(t)}$, where $\lambda\phi(t)$ is the cumulative hazard rate and we interpret $\phi(t) = \int_0^t F(p(s))ds$ as the cumulative innovation incentive offered by the social planner to date. Unlike in Section 3, the environment is not stationary, so $p(t)$ and the rate of growth in $\phi(t)$ (given by (2)) will change over time.

Applying Bayes' Rule, the social planner's beliefs at time t are

$$h_t(\lambda) = \frac{h_0(\lambda)e^{-\lambda\phi(t)}}{\int_0^\infty h_0(\lambda')e^{-\lambda'\phi(t)}d\lambda'} \tag{11}$$

An important implication of the updating formula (11) is that the social planner's beliefs depend only on the level of ϕ and not on how much time it takes to reach ϕ nor on other aspects of the history of the reward policy. That is, the parameter ϕ alone is a sufficient statistic for the information that is relevant to the social planner's optimization problem. As a result, we can model the optimization problem as a Markov process where the value function and the optimal policy are functions of ϕ.

Although the reward policy does not appear in the updating formula (11), the function ϕ incorporates the impact of the reward on the social planner's learning process. If the social planner is conservative with the policy and sets p at a low level until time t, then equations (1) and (2) show that ϕ increases slowly and will be relatively low. In this case, we can see from that (11) that the updated belief will be relatively close to $h_0(\lambda)$.

To take advantage of the Markov structure of our model, we will use ϕ rather than calendar time t as the state variable.[15] At state ϕ, the social planner's beliefs about the arrival rate λ are

$$h_\phi(\lambda) = \frac{h_0(\lambda)e^{-\lambda\phi}}{\int_0^\infty h_0(\lambda')e^{-\lambda'\phi}d\lambda'} \tag{12}$$

[15] So long as $\frac{d\phi}{dt} > 0$, the results in our model can be translated to calendar time using the change of variables $dt = \frac{d\phi}{F(p(t))}$.

and the cumulative distribution is $H_\phi(\lambda) = \int_0^\lambda h_\phi(\lambda')d\lambda'$. Equation (12) implies that $h_\phi(\lambda)$ is continuous in ϕ.

Our first result characterizes how the social planner's beliefs change as the cumulative innovation incentive ϕ increases, using the concept of stochastic dominance. Lemma 1 formalizes the intuitive result that as ϕ increases, the social planner becomes more pessimistic.

Lemma 1 *[Increasing pessimism] If $\phi_1 < \phi_2$, the distribution H_{ϕ_1} dominates the distribution H_{ϕ_2} in the sense of first order stochastic dominance.*

The lemma follows directly from the fact that the social planner is only able to observe an idea if the recipient of the idea invests in it. When the reward is not claimed, the signal about the arrival rate is negative, so the planner becomes more pessimistic.

A follow-up question is how quickly the social planner gets pessimistic over time. Defining this will be crucial for the characterization of the dynamics of the optimal policy. For this purpose, we use the moments of $h_\phi(\lambda)$ which are given by

$$m_k(\phi) = \int_0^\infty \lambda^k h_\phi(\lambda)d\lambda \text{ for } k = 1, 2, ... \tag{13}$$

In particular, the expected value of λ is $m_1(\phi)$. Because $h_\phi(\lambda)$ is continuous in ϕ, equation (13) implies that $m_k(\phi)$ are continuous in ϕ also.

Our next lemma presents differential equations for the evolution of the moments of the social planner's beliefs $h_\phi(\lambda)$. These dynamics are consistent with the pessimism result of Lemma 1 in that all the moments of the social planner's beliefs are decreasing in ϕ. In particular, the expected value of λ (the first moment) is decreasing. However, Lemma 2 goes beyond Lemma 1 in that it explicitly characterizes the rates at which the moments evolve.

Lemma 2 *[Evolution of the moments] For $k = 1, 2, ...$*

(i) The moments $m_k(\phi)$ of the social planner's beliefs $h_\phi(\lambda)$ are finite, continuously differentiable, and strictly positive for all ϕ. The moments are strictly decreasing in ϕ with

derivatives given by

$$\frac{dm_k(\phi)}{d\phi} = m_k(\phi)m_1(\phi) - m_{k+1}(\phi) = -Cov_\phi(\lambda, \lambda^k) < 0 \tag{14}$$

In particular, the derivative of the expected value of λ is

$$\frac{dm_1(\phi)}{d\phi} = m_1^2(\phi) - m_2(\phi) = -Var_\phi(\lambda) < 0 \tag{15}$$

(ii) As $\phi \to \infty$, $m_k(\phi) = \frac{k!}{\phi^k} + o(\frac{1}{\phi^k})$ so that $lim_{\phi\to\infty} m_k(\phi) = 0$.

(iii) As $\phi \to \infty$, the ratio $\frac{m_{k+1}(\phi)}{m_k(\phi)}$ is strictly decreasing with $\frac{m_{k+1}(\phi)}{m_k(\phi)} = \frac{k+1}{\phi} + o(\frac{1}{\phi})$ so that $lim_{\phi\to\infty} \frac{m_{k+1}(\phi)}{m_k(\phi)} = 0$.

Lemma 2 provides a full characterization of the social planner's beliefs as a function of ϕ.[16] Notice that the equations of motion (14) and (15) are independent of the social planner's reward policy. However, these dynamics are an essential ingredient to the optimal reward, because the level of the reward determines how quickly the state ϕ increases (as shown in (2)) and hence how fast the beliefs ultimately evolve.

4.2 Optimal policy

We are now ready to formulate the social planner's optimization problem. At time $t = 0$, the state variable is $\phi(0) = 0$, and the value function is

$$V(0) = \max_{p(\cdot)} \int_0^\infty \int_0^\infty e^{-(rt+\lambda\phi(t))} \left(\frac{v}{r} - E(c|p(t))\right) \lambda F(p(t)) h_{\phi(t)}(\lambda) d\lambda dt \tag{16}$$
$$\text{subject to } \frac{d\phi(t)}{dt} = F(p(t))$$

As in our benchmark model, the social planner faces a trade-off between the cost of innovation and delay. A higher reward policy $p(t)$ at time t would imply less delay, since the arrival rate of viable ideas $\lambda F(p(t))$ is higher, but it would also imply a higher expected cost of innovation $E(c|p(t))$. However, the optimization problem now incorporates the evolution

[16] Lemma 2 shows how the evolution of a probability distribution may be characterized through the dynamics of the moments. This approach may prove useful in other contexts. If the social planner's prior h_0 is uniquely determined by its moments, then h_ϕ is as well. In this case, the vector of moments $(m_k(\phi))$ fully characterizes h_ϕ without referencing to h_0. A sufficient condition for this is that h_0 satisfies Carleman's condition.

of the planner's beliefs $h_{\phi(t)}$. To solve the model, we again apply standard techniques from dynamic programming. We approximate continuous time with a discrete time interval Δt, solve that problem and then take the limit of the solution (the level of social welfare) as $\Delta t \to 0$. The discount rate $e^{-r\Delta t}$ is again approximated by $(1 - r\Delta t)$ and the probability that an idea arrives is $\lambda \Delta t$. We use Bellman's principle of optimality to express the social planner's optimization problem recursively as the sum of a current payoff and a future continuation payoff as follows:

$$V(\phi) = \max_{p \in [\underline{c}, \overline{c}]} \left\{ \begin{array}{c} m_1(\phi)\Delta t \int_{\underline{c}}^{p} \left(\frac{v}{r} - c\right) f(c)dc + \\ (1 - m_1(\phi)\Delta t F(p))\,(1 - r\Delta t)V(\phi + \Delta\phi) \end{array} \right\} \tag{17}$$

where

$$\Delta\phi = F(p)\Delta t \tag{18}$$

Here, (18) is the discrete time analogue of (2) above. The first term on the right-hand side of (17) is the social welfare if a viable idea arrives and is invested. The probability of this outcome is $m_1(\phi)\Delta t F(p)$ since $m_1(\phi)$ is the expected value of λ and $F(p)$ is the probability that the idea has a cost less than the reward. The second term is the social welfare to be realized in the future if no idea is invested today. The probability of this outcome is $1 - m_1(\phi)\Delta t F(p)$. The future payoff $V(\phi + \Delta\phi)$ is discounted by $(1 - r\Delta t)$.

In (17), we assume that $p(t) \in [\underline{c}, \overline{c}]$, which is the support of the cost distribution F. A policy $p < \underline{c}$ cannot be optimal at any time t because it will not incentivize investment, and the social planner also would not learn anything ($\Delta\phi = 0$). Such a policy merely results in lost time, as the optimization problem will be the same in the next period as it is today. Any policy $p > \overline{c}$ yields the same welfare as $p = \overline{c}$, as there are no ideas with cost $c > \overline{c}$, and the social planner learns the maximal amount possible under all such policies ($\Delta\phi = \Delta t$). Hence, there is no loss in generality in assuming that $p(t) \leq \overline{c}$. We will show that the optimal policy is below $\frac{v}{r}$, as in the benchmark model. However, before doing this, we must account for the influence of learning on the choice of the optimal policy.

The key difference between the Bellman equation (17) in this section and (4) in the last section is that the value function V in (17) changes as the social planner's beliefs change.

In choosing a reward policy, the social planner knows that the higher the reward is today, the more pessimistic beliefs will be tomorrow if the reward is not claimed.

Subtracting $V(\phi)$ from both sides of equation (17), dividing by Δt, and passing to the limit $\Delta t \to 0$, we obtain the Hamilton-Jacobi-Bellman equation for the social planner's optimization problem in continuous time:

$$0 = \max_{p \in [\underline{c}, \bar{c}]} \left\{ m_1(\phi) \int_{\underline{c}}^{p} \left(\frac{v}{r} - c \right) f(c) dc - (r + m_1(\phi) F(p)) V(\phi) + \frac{dV}{d\phi} \frac{d\phi}{dt} \right\} \tag{19}$$

where $\frac{d\phi}{dt} = F(p)$. Taking the derivative of the right-hand side of equation (19) with respect to p, the first order condition for the optimal reward policy is

$$0 = m_1(\phi) f(p(\phi)) \left(\frac{v}{r} - p(\phi) - V(\phi) \right) + f(p(\phi)) \frac{dV(\phi)}{d\phi} \tag{20}$$

which simplifies to

$$V(\phi) = \left(\frac{v}{r} - p(\phi) \right) + \frac{1}{m_1(\phi)} \frac{dV(\phi)}{d\phi} \ . \tag{21}$$

In the appendix, we derive a relationship between the value function and the optimal policy:

$$V(\phi) = \frac{m_1(\phi)}{r} \int_{\underline{c}}^{p(\phi)} F(c) dc \tag{22}$$

This relationship is analogous to equation (9) for the static model, with the expected value of the arrival rate $m_1(\phi)$ replacing λ.

From (22), the derivative of $V(\phi)$ is

$$\frac{dV(\phi)}{d\phi} = \frac{1}{r} \left(\frac{dm_1(\phi)}{d\phi} \int_{\underline{c}}^{p(\phi)} F(c) dc + m_1(\phi) F(p(\phi)) \frac{dp(\phi)}{d\phi} \right) \tag{23}$$

In the appendix, we use equation (23) and Lemma 2 to derive the differential equation for the optimal policy:[17]

$$\frac{dp(\phi)}{d\phi} = \frac{1}{F(p(\phi))} \left(\frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p(\phi)} F(c) dc - r \left(\frac{v}{r} - p(\phi) \right) \right) \tag{24}$$

[17]By transforming the variable ϕ to t using $\frac{d(\phi(t))}{dt} = F(p)$, we can also express equation (24) as a function of time: $\frac{dp(t)}{dt} = \frac{m_2(\phi(t))}{m_1(\phi(t))} \int_{\underline{c}}^{p(t)} F(c) dc - r(\frac{v}{r} - p(t)))$.

Theorem 3 *The optimal reward policy $p(\phi)$ is the unique solution to (22). As ϕ increases, $p(\phi)$ increases and social welfare $V(\phi)$ decreases. As $\phi \to \infty$, the optimal reward policy increases to $\frac{v}{r}$ at the rate $p(\phi) = \frac{v}{r} - \frac{2}{\phi r} \int_{\underline{c}}^{\frac{v}{r}} F(c)dc + o(\frac{1}{\phi})$.*

Intuitively, $V(\phi)$ is decreasing because the social planner becomes more and more pessimistic about the arrival rate of ideas as time continues without a viable idea. To mitigate delay, it is optimal to tolerate higher cost. However, the social planner now also takes learning into account. Comparing the first order condition (21) to the first order condition (6) in the benchmark model shows that the first order condition now has an additional term accounting for how V changes as ϕ increases. Since V is decreasing, the learning term is negative implying that social welfare is lower than the value of the marginal idea, instead of being equal to it as in the benchmark model. This implies that there will be ideas which will not be invested in even though $v/r - c$ is greater than $V(\phi)$. This wedge reflects the social planner's incentive to learn about the arrival rate of ideas.

One way to assess the role of learning is to compare the optimal policy $p(\phi)$ to the policy that a social planner would choose if λ were fixed at the expected value $m_1(\phi)$. That is, we consider a naive social planner who ignores learning and assumes that $\lambda = m_1(\phi)$. We ask how the reward that this social planner would choose compares to the dynamically optimal reward $p(\phi)$ defined by (22).

Theorem 4 *The optimal reward policy $p(\phi)$ is lower than it would be if the arrival rate were fixed at $\lambda = m_1(\phi)$.*

Theorem 4 shows that the uncertainty over λ causes the social planner to choose a more conservative reward policy. Because the social planner and innovators are risk neutral, this result does not reflect risk aversion. Instead, it reflects the nature of learning and the option value of delay. Information has value because it allows society to achieve a better tradeoff between investment and delay. However, information is not useful once investment happens. This effect encourages the social planner to be more conservative: the lower policy prevents high-cost ideas from being invested in while allowing society to gather additional information about the scarcity of ideas.[18]

Our next result provides a lower bound for the optimal policy. We compare the optimal policy to the policy that a social planner would have chosen if λ were fixed at the social planner's expected arrival rate in the moment after observing an innovation.

Theorem 5 *The optimal reward policy $p(\phi)$ is higher than it would be if the arrival rate were fixed at $\lambda = \frac{m_2(\phi)}{m_1(\phi)}$.*

In the moment after innovation, when the reward is claimed, the social planner becomes more optimistic about the arrival rate. This being the best "news" possible, it is intuitive that a social planner facing this arrival rate would choose a lower reward. In the proof of Theorem 5, we derive the expression $\lambda = \frac{m_2(\phi)}{m_1(\phi)}$ using the martingale property of Bayesian updating.

4.3 Scarcity of low cost ideas

Our last result provides a comparative static for the cost distribution of ideas. Recall that in the benchmark model, we showed that a shift in weight from lower to higher costs in the sense of first order stochastic dominance reduces social welfare and causes the social planner to set a more generous reward policy (Theorem 2). Although this result is very intuitive (higher costs should imply lower welfare), we need a stronger form of stochastic dominance to ensure a similar comparative static in the model with learning. We again compare the cost distribution F to a second cost distribution G with support $[\underline{c}, \overline{c}]$ and a continuous probability density $g > 0$ on the support.

Definition 1 *[Strict Monotone Likelihood Ratio Property] The cost distribution F stochastically dominates the cost distribution G in the sense of the strict monotone likelihood ratio property if the ratio $f(c)/g(c)$ is strictly increasing in c for $c \in [\underline{c}, \overline{c}]$.*

[18] In an online appendix, we analyze a full information environment where the social planner observes the arrival of ideas. As time passes, the social planner updates its beliefs based on the number of ideas k that arrive in time t. In this model, the reward policy does not impact the evolution of the social planner's beliefs. As a result the social planner has no incentive to delay innovation to learn more about the arrival rate. We show that Theorem 4 no longer holds.

Under the strict monotone likelihood ratio property, the relative likelihood of a particular cost arising under f than under g is higher for higher costs.

Theorem 6 *If the cost distribution F stochastically dominates the cost distribution G in the sense of the strict monotone likelihood ratio property, then social welfare is lower under F than under G and the optimal reward policy is higher.*

The simplest intuition for this comparative static is the same as in our benchmark model: for any level of the reward policy, fewer ideas are viable under F than under G, so the probability of investment is lower. This can explain both why social welfare is lower and why the optimal reward policy is higher under F. However, the fact that the social planner is learning about the scarcity parameter λ brings an additional effect into play that acts in the opposite direction. Because fewer ideas are viable under F, the fact that a prize is not claimed provides a less negative signal, so the social planner does not grow pessimistic as quickly. This signalling effect leads the planner to increase the prize more slowly under F than under G. As shown in Theorem 6, under the strict monotone likelihood ratio property we assume, the first effect dominates and drives the comparative statics.

5 Conclusion

Optimal decision making under scarcity is central to economics. Our goal in this paper is to introduce a framework where ideas have economic value because they are scarce. We consider how the scarcity of ideas for innovation affects the social planner's prize structure. Scarcity of ideas means that if the social planner fails to invest in an idea, the next idea may not arrive for a while.

When ideas arrive to random agents at random times, we have argued that the fundamental trade-off faced by a social planner determining the optimal incentive structure is between cost and delay. The speed with which innovations emerge in the economy can be faster, but it would come at a higher cost. Our findings reveal that rewards should be higher in environments where ideas are scarce. If ideas are scarce, higher cost should be tolerated in order to reduce delay.

We model the scarcity of ideas as a Poisson arrival rate. In our benchmark model, the social planner knows the true value of the arrival rate. In our main model, the social planner is uncertain about the arrival rate and updates her prior beliefs as time passes without an innovation. We find that the cumulative hazard function for the arrival process provides a one-dimensional Markov state variable that allows us to derive differential equations for how the beliefs and the reward policy evolve over time.

We show that in this dynamic model, the optimal reward should be increasing as time passes without an innovation. The result is driven by the fact that the social planner learns about scarcity over time. Longer delay reveals that ideas are more scarce, and so it leads to expectation of an even longer delay. The delay can be mitigated with higher rewards since higher rewards encourage investment in higher-cost ideas. We show that when the arrival rate is not known, the social planner sets a lower reward than if the arrival rate were known to be equal to the expected value of beliefs in the dynamic model. This leads to more delay as society learns more about the true value of the scarcity.

In our framework, different ideas are substitutes for filling the same market niche and are distinguished from each other by their costs. This raises the question of how the reward policies in our two models depend on the distribution of costs. For each model, we identify stochastic dominance assumptions such that reward policies are more generous if low-cost ideas are expected to be more scarce.

Our framework brings to light an additional role for the R&D reward structure. It should ensure that idea recipients only invest if the value of the investment outweighs the option value of waiting for a better idea. Our results apply equally well to patents and prizes, and any other way of giving rewards.[19] Patents raise the issue of whether our prescriptions can be implemented under existing patent doctrine. They also raise the question of how deadweight loss incurred in collecting the reward money changes the optimal policy. For example, in the United States, the main requirements for obtaining a patent are novelty,

[19]For a sample of the many ways, other than patents, that economists have thought about incentives in R&D, see Wright (1983), Chapters 2 and 8 of Scotchmer (2004), and Hopenhayn, Llobet and Mitchell (2006).

non-obviousness, utility and enablement.[20] Together, these requirements govern the breadth of claims that are granted. When the statutory patent life is the same for all patentable innovations, breadth is the main lever to differentiate rewards. Our prescription is therefore that patent offices and courts should grant generous claims (broad patents) when ideas are scarce, or more particularly, when the innovation arrives after long delay.

Patent law doctrine also has a threshold standard for granting a patent, namely, the non-obviousness requirement. Our arguments can be interpreted to mean that this threshold standard should be interpreted more leniently when ideas are scarce. In fact, patent doctrine has its own term for this circumstance, namely, "long-felt need." Long-felt need is one of the secondary considerations for patentability. Long-felt need is a criterion that can be used to fulfill the non-obviousness requirement of a patent application. To meet the requirement, an innovation must address a problem that is recognized for a long time without a solution by "those of ordinary skill in the art."[21]

For future research, using the notion of substitute ideas, we provide a framework of thinking about the scarcity of ideas. Our model of ideas, and the fact that ideas are private, can be interpreted as a model of creativity or imagination. Oddly, it is hard to find creativity in the economists' tool kit for studying R&D. Our approach suggests a direction of research that might fill that gap. Because ideas are not common knowledge in our model, innovators make positive profit in expectation. In environments where ideas are scarce, only the very creative receive ideas and make profit.

In this paper, we considered the selection process of a social planner who has the goal of maximizing social welfare. Our modelling approach can also be adapted to study the selection process which takes place in the R&D departments of private firms, where managers maximize profits (instead of social welfare) by using a gate system to decide which ideas go further and which are eliminated. In such laboratory environments, if qualified researchers have tried for a long time to solve a problem without success, the limiting resource is idea scarcity rather than R&D facilities.

[20] See https://www.uspto.gov/web/offices/pac/mpep/consolidated_laws.pdf.

[21] See https://www.uspto.gov/web/offices/pac/mpep/s716.html.

Future work can also build on our framework by adding competition to the model. The impact of competition is purposefully shut down in our model because our purpose is to parameterize the scarcity of ideas and to study how a social planner should formulate policy to select between substitute ideas. While our framework allows us to capture the notions of nonobviousness and long-felt need that are used in patent law, competition is restricted in our model by the exogenous process which governs the arrival of ideas. Moving forward, the insights from our model can be enriched by considering competition in a number of ways. One way of introducing competitive pressure is by allowing innovators to bank their ideas. If innovators can bank their ideas to claim the reward at a future date, they may be pre-empted by other idea recipients. This puts competitive pressure on their decisions. Another way of introducing competition to the model may be by assuming that the arrival rate of ideas is determined by both exogenous and endogenous factors. Yet a third way is to consider imitation of ideas.

Finally, we considered an environment where ideas are heterogeneous in terms of their cost of implementation. Considering an environment where ideas are heterogeneous in terms of their values would also be interesting.[22] In such a model, there may be multiple complementary market niches, or filling one market niche may open further opportunities.

[22] In this case, it would make sense for the social planner to reward innovators with patents. If the social planner used a prize system, then there would be no way of preventing ideas with low values from being invested in. Given a prize, an innovator would invest if and only if the prize value exceeds the cost irrespective of the value of the idea. If the social planner uses a patent system with duration T, then innovators will invest if and only if $v/r(1 - e^{-rT}) \geq c$.

References

[1] Agrawal, A., J. S. Gans, and S. Stern, 2021. "Enabling entrepreneurial choice," *Management Science*, 67(9), 5510-5524.

[2] Achdou, Y., J. Han, J. Lasry, P. Lions, and B. Moll. 2022. "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *Review of Economics Studies,* 89, 45-86.

[3] Bryan, K. A., and J. Lemus. 2017. "The Direction of Innovation," *Journal of Economic Theory*, 172, 247-272.

[4] Cao, D. and W. Luo. 2017. "Persistent Heterogeneous Returns and Top End Wealth Inequality," *Review of Economic Dynamics,* 26, 301-326.

[5] Choi, J. P. 1991. "Dynamic R&D Competition under 'Hazard Rate' Uncertainty," *RAND Journal of Economics,* 22, 596-610

[6] Dixit, A. and R. Pindyck. 1994. *Investment under Uncertainty.* Princeton, NJ: Princeton University Press.

[7] Erkal, N. and S. Scotchmer. 2007. "Scarcity of Ideas and Options to Invest in R&D," University of California, Berkeley, Department of Economics, Working Paper 07-348.

[8] Erkal, N. and S. Scotchmer. 2009. "Scarcity of Ideas and R&D Options: Use it, Lose it, or Bank it," NBER Working Paper No. 14940.

[9] Green, J. R., and S. Scotchmer. 1995. "On the division of profit in sequential innovation," *RAND Journal of Economics*, 26, 20-33.

[10] Hopenhayn, H., G. Llobet, and M. Mitchell. 2006. "Rewarding Sequential Innovators: Patents, Prizes, and Buyouts." *Journal of Political Economy,* 114, 1041-1068.

[11] Kohn, M. G. and S. Shavell. 1974. "The Theory of Search." *Journal of Economic Theory,* 9(2), 93-123.

[12] Lee, T. and L. L. Wilde. 1980. "Market structure and innovation: A reformulation." *Quarterly Journal of Economics,* 94(2), 429-436.

[13] Loury, G. C. 1979. "Market structure and innovation." *Quarterly Journal of Economics,* 93(3), 395-410.

[14] McCall, B. P. and J. J. McCall. 2008. *Economics of Search.* New York: Routledge Press.

[15] MacDonald, R. and D. Siegel. 1986. "The Value of Waiting to Invest," *Quarterly Journal of Economics*, 101, 707-728.

[16] Malueg, D. A. and S. O. Tsutsui. 1997. "Dynamic R&D Competition with Learning," *RAND Journal of Economics,* 28, 751-772.

[17] Mason, R. and J. Välimäki, 2011. "Learning About the Arrival of Sales," *Journal of Economic Theory,* 145(4), 1699-1711.

[18] Nuno, G. and B. Moll. 2018. "Social Optima in Economies with Heterogeneous Agents," *Review of Economic Dynamics,* 28, 150-180.

[19] O'Donoghue, T., Scotchmer, S. and Thisse, J.-F. 1998. "Patent Breadth, Patent Life and the Pace of Technological Progress," *Journal of Economics and Management Strategy*, 7, 1-32.

[20] Rosenfield, D. B. and R. D. Shapiro. 1981. "Optimal Adaptive Price Search," *Journal of Economic Theory*, 25, 1-20.

[21] Rothschild, M. 1974. "Searching for the Lowest Price when the Distribution of Prices is Unknown," *Journal of Political Economy*, 82, 689-711.

[22] Scotchmer, S. 2004. *Innovation and Incentives.*Cambridge, MA: MIT Press.

[23] Stokey, N. L., R. E. Lucas and E. C. Prescott. 1989. *Recursive Methods in Economic Dynamics.* Cambridge, MA: Harvard University Press.

[24] Weeds, H. 2002. "Strategic Delay in a Real Options Model of R&D Competition," *Review of Economic Studies*, 69, 729-747.

[25] Wright, B. D. 1983. "The Economics of Invention Incentives: Patents, Prizes and Research Contracts," *American Economic Review,* 73, 691-707.

APPENDIX

1 Proofs for Section 3

Proof of Theorem 1. Equation (4) is the Bellman equation for the discrete time version of the problem. The maximal value of social welfare is $\frac{v}{r} - \underline{c}$ and the minimum value is 0. Let S denote the compact interval $[0, \frac{v}{r} - \underline{c}]$. We define the Bellman operator as a function from S to S as follows:

$$TW = \max_{p} \left\{ \lambda \Delta t \int_{\underline{c}}^{p} \left(\frac{v}{r} - c \right) f(c)dc + (1 - \lambda \Delta t F(p)) (1 - r\Delta t) W \right\} \tag{25}$$

where $W \in S$. It is straightforward to check that the Bellman operator T is a contraction mapping on S. From standard theory, it has a unique fixed point in S.[23] We denote this fixed point $V^{\Delta t}$. The unique solution to the continuous time version of the problem (equation (4)) is then $V = \lim_{\Delta t \to 0} V^{\Delta t}$. The optimal reward policy p^* is the unique solution to equation (9) for this value of V.

To prove the comparative static results for λ, we write the optimal policy p^* as p^*_λ. From (8), we have

$$\frac{v}{r} - p^*_\lambda = \frac{\lambda}{r} \int_{\underline{c}}^{p^*_\lambda} F(c)dc. \tag{26}$$

Differentiating equation (26) with respect to λ we have

$$-\frac{dp^*_\lambda}{d\lambda} = \frac{1}{r} \int_{\underline{c}}^{p^*_\lambda} F(c)dc + \frac{\lambda}{r} F(p^*_\lambda) \frac{dp^*_\lambda}{d\lambda}$$

Rearranging, we have

$$\frac{dp^*_\lambda}{d\lambda} = -\frac{\int_{\underline{c}}^{p^*_\lambda} F(c)dc}{r + \lambda F(p^*_\lambda)} < 0$$

Thus, the optimal policy p^*_λ is decreasing in λ. The result that social welfare V_λ is increasing in λ now follows immediately from equation (6). ■

Proof of Theorem 2. Let p^*_F and p^*_G be the optimal policies under the cost distributions F and G respectively. Let V_F and V_G denote the corresponding value functions. From

[23]See Stokey et al. (1989), Chapter 3, in particular Theorems 3.2 (Contraction Mapping Theorem), 3.3 (Blackwell's sufficient conditions for a contraction), and 3.6 (Theorem of the Maximum).

equations (26) and (6), we have

$$V_F = \frac{v}{r} - p_F^* = \frac{\lambda}{r}\int_{\underline{c}}^{p_F^*} F(c)dc \tag{27}$$

$$V_G = \frac{v}{r} - p_G^* = \frac{\lambda}{r}\int_{\underline{c}}^{p_G^*} G(c)dc \tag{28}$$

Since F stochastically dominates G, we have $F(c) \leq G(c)$ for all c and hence that $\int_{\underline{c}}^{x} F(c)dc \leq \int_{\underline{c}}^{x} G(c)dc$ for all x. Thus (27) implies

$$\frac{v}{r} - p_F^* \leq \frac{\lambda}{r}\int_{\underline{c}}^{p_F^*} G(c)dc \tag{29}$$

But this implies that the policy that makes (29) hold with equality is lower than p_F^*. (The right hand side of (29) is increasing in p and the left hand side is decreasing in p.) That is: $p_G^* \leq p_F^*$. That social welfare is lower under F than G now follows from (27) and (28). ■

2 Proofs for Section 4

Proof of Lemma 1. The beliefs $h_\phi(\lambda)$ are given in equation (12). Taking the derivative with respect to ϕ we have[24]

$$\begin{aligned}\frac{dh_\phi(\lambda)}{d\phi} &= \frac{(-\lambda h_0(\lambda)e^{-\lambda\phi})}{\int_0^\infty h_0(\lambda')e^{-\lambda'\phi}d\lambda'} + \frac{h_0(\lambda)e^{-\lambda\phi}}{\int_0^\infty h_0(\lambda')e^{-\lambda'\phi}d\lambda'}\frac{\int_0^\infty \lambda' h_0(\lambda')e^{-\lambda'\phi}d\lambda'}{\int_0^\infty h_0(\lambda')e^{-\lambda'\phi}d\lambda'} \\ &= -\lambda h_\phi(\lambda) + h_\phi(\lambda)m_1(\phi) \\ &= h_\phi(\lambda)(m_1(\phi) - \lambda) \end{aligned} \tag{30}$$

where $m_1(\phi)$ is the expected value of h_ϕ. Since $h_\phi(\lambda) > 0$ for all $\lambda \in R_+$, equation (30) implies that $\frac{dh_\phi(\lambda)}{d\phi}$ is single-peaked. That is: $\frac{dh_\phi(\lambda)}{d\phi} > 0$ for $\lambda < m_1(\phi)$ and $\frac{dh_\phi(\lambda)}{d\phi} < 0$ for $\lambda > m_1(\phi)$.

To prove the lemma, it suffices to show that $\frac{d}{d\phi}H_\phi(\lambda) \geq 0$ for $\lambda \geq 0$ and $\phi \geq 0$. For any ϕ, consider $\lambda \leq m_1(\phi)$. From (30), we have

$$\frac{d}{d\phi}H_\phi(\lambda) = \frac{d}{d\phi}\int_0^\lambda h_\phi(\lambda')d\lambda' = \int_0^\lambda \frac{dh_\phi(\lambda')}{d\phi}d\lambda' \geq 0$$

[24]Note that we are taking derivatives inside integral signs. This is possible because $\lambda h_0(\lambda)e^{-\lambda\phi}$ is dominated by $\lambda h_0(\lambda)$ which is integrable given Assumption 1 which states that the moments of h_0 are finite.

For $\lambda > m_1(\phi)$, we have

$$\begin{aligned}\frac{d}{d\phi}H_\phi(\lambda) &= \frac{d}{d\phi}\int_0^\lambda h_\phi(\lambda')d\lambda' = \frac{d}{d\phi}\int_0^\infty h_\phi(\lambda')d\lambda' - \frac{d}{d\phi}\int_\lambda^\infty h_\phi(\lambda')d\lambda' \\ &= 0 - \frac{d}{d\phi}\int_\lambda^\infty h_\phi(\lambda')d\lambda' = -\int_\lambda^\infty \frac{dh_\phi(\lambda')}{d\phi}d\lambda' > 0\end{aligned}$$

where we have used the facts that $\int_0^\infty h_\phi(\lambda')d\lambda' = 1$ (so its derivative is 0) and that $\frac{dh_\phi(\lambda)}{d\phi} < 0$ for $\lambda > m_1(\phi)$. ■

Proof of Lemma 2.

Part 1. In this part, we show that the moments $m_k(\phi)$ are finite and strictly positive. We also calculate their derivatives and show that these are continuous.

By Assumption 1, the moments $m_k(0)$ of the initial beliefs $h_0(\lambda)$ are finite for $k = 1, 2, ...$ From (11), we have

$$m_k(\phi) = \frac{\int_0^\infty \lambda^k h_0(\lambda)e^{-\lambda\phi}d\lambda}{\int_0^\infty h_0(\lambda')e^{-\lambda'\phi}d\lambda'} \tag{31}$$

That $m_k(\phi)$ is finite for $\phi > 0$ follows from (31), the fact that $e^{-\lambda\phi} < 1$, and the fact that $\lambda^k h_0(\lambda)$ and $h_0(\lambda)$ have finite integrals (equal to $m_k(0)$ and 1 respectively). That $m_k(\phi) > 0$ for $\phi > 0$ follows immediately from (31).

Taking the derivative of (31) using the quotient rule,[25] we have

$$\begin{aligned}\frac{dm_k(\phi)}{d\phi} &= \frac{(\int_0^\infty \lambda^k h_0(\lambda)e^{-\lambda\phi}d\lambda)(\int_0^\infty \lambda h_0(\lambda)e^{-\lambda\phi}d\lambda)}{(\int_0^\infty h_0(\lambda)e^{-\lambda\phi}d\lambda)^2} - \frac{\int_0^\infty \lambda^{k+1}h_0(\lambda)e^{-\lambda\phi}d\lambda}{\int_0^\infty h_0(\lambda)e^{-\lambda\phi}d\lambda} &(32) \\ &= m_k(\phi)m_1(\phi) - m_{k+1}(\phi) = -Cov_\phi(\lambda, \lambda^k) < 0 &(33)\end{aligned}$$

The covariance, $Cov_\phi(\lambda, \lambda^k)$, is strictly positive because $h_\phi(\lambda)$ is a non-degenerate density and λ^k is an increasing function of λ. For $k = 1$,

$$\frac{dm_1(\phi)}{d\phi} = m_1(\phi)m_1(\phi) - m_2(\phi) = -Cov_\phi(\lambda, \lambda) = -Var_\phi(\lambda) < 0$$

Since $m_k(\phi)$ is continuous for $k = 1, 2, ..$, equation (33) shows that the $\frac{dm_k(\phi)}{d\phi}$ are continuous.

[25] We can differentiate under the integrals in (31) because each integrand that appears on the right-hand side of (32) is dominated by either $h_0(\lambda)$ or by $\lambda^k h_0(\lambda)$ for some k, and these functions have finite integrals.

Part 2. In this part, we show that $\lim_{\phi\to\infty} m_k(\phi) = 0$, and that $m_k(\phi) = \frac{k!}{\phi^k} + o(\frac{1}{\phi^k})$ as s $\phi \to \infty$. We use a change of variable, $z = \lambda\phi$, to express $\lim_{\phi\to\infty} m_k(\phi)$ as

$$\lim_{\phi\to\infty} m_k(\phi) = \lim_{\phi\to\infty} \frac{\int_0^\infty \lambda^k h_0(\lambda) e^{-\lambda\phi} d\lambda}{\int_0^\infty h_0(\lambda) e^{-\lambda\phi} d\lambda} = \lim_{\phi\to\infty} \frac{1}{\phi^k} \frac{\int_0^\infty z^k h_0(\frac{z}{\phi}) e^{-z} dz}{\int_0^\infty h_0(\frac{z}{\phi}) e^{-z} dz}$$

Taking the limit inside the integrals,[26] we have

$$\lim_{\phi\to\infty} m_k(\phi) = \lim_{\phi\to\infty} \frac{1}{\phi^k} \frac{\int_0^\infty z^k h_0(0) e^{-z} dz}{\int_0^\infty h_0(0) e^{-z} dz} = \lim_{\phi\to\infty} \frac{\Gamma(k+1)}{\phi^k} \tag{34}$$

where $\Gamma(x) = \int_0^\infty z^{x-1} e^{-z} dz$ is the gamma function. For positive integers k the gamma function has the value $\Gamma(k) = (k-1)!$ so we have now shown that

$$\lim_{\phi\to\infty} m_k(\phi) = \lim_{\phi\to\infty} \frac{k!}{\phi^k}$$

This immediately implies that $\lim_{\phi\to\infty} m_k(\phi) = 0$ and that $m_k(\phi) = \frac{k!}{\phi^k} + o(\frac{1}{\phi^k})$.

Part 3. In this part, we show that $\frac{m_{k+1}(\phi)}{m_k(\phi)}$ is strictly decreasing in ϕ and that $\lim_{\phi\to\infty} \frac{m_{k+1}(\phi)}{m_k(\phi)} = 0$ with $\frac{m_{k+1}(\phi)}{m_k(\phi)} = \frac{k+1}{\phi} + o(\frac{1}{\phi})$.

We first show that $\frac{m_2(\phi)}{m_1(\phi)}$ is strictly decreasing. Using the expressions for $\frac{dm_1(\phi)}{d\phi}$ and $\frac{dm_2(\phi)}{d\phi}$ in (33), we have

$$\frac{d}{d\phi}\left(\frac{m_2(\phi)}{m_1(\phi)}\right) = \frac{m_2^2(\phi) - m_1(\phi) m_3(\phi)}{m_1^2(\phi)}$$

Therefore, $\frac{d}{d\phi}\frac{m_2(\phi)}{m_1(\phi)} < 0$ if and only if

$$m_2^2(\phi) < m_1(\phi) m_3(\phi) \tag{35}$$

Using the definition of $m_k(\phi)$ in (31), it is straightforward to show that the inequality (35) holds if and only if

$$\left(\int_0^\infty \lambda^2 h_0(\lambda) e^{-\lambda\phi} d\lambda\right)^2 < \int_0^\infty \lambda h_0(\lambda) e^{-\lambda\phi} d\lambda \int_0^\infty \lambda^3 h_0(\lambda) e^{-\lambda\phi} d\lambda. \tag{36}$$

We can write this as

$$\left(\int_0^\infty g_1(\lambda) g_2(\lambda) d\lambda\right)^2 < \int_0^\infty g_1^2(\lambda) d\lambda \int_0^\infty g_2^2(\lambda) d\lambda \tag{37}$$

[26]We can bring the limit inside the integral signs because of Assumption 1.

where $g_1(\lambda) = (\lambda h_0(\lambda) e^{-\lambda\phi})^{1/2}$ and $g_2(\lambda) = (\lambda^3 h_0(\lambda) e^{-\lambda\phi})^{1/2}$ are positive square integrable functions. The inequality then follows from the Cauchy-Schwartz inequality for positive square-integrable functions.

A similar argument shows that $\frac{m_{k+1}(\phi)}{m_k(\phi)}$ is strictly decreasing in ϕ. We again use the expressions for $\frac{dm_{k+1}(\phi)}{d\phi}$ and $\frac{dm_k(\phi)}{d\phi}$ in (33), and we take $g_1(\lambda) = (\lambda^k h_0(\lambda) e^{-\lambda\phi})^{1/2}$ and $g_2(\lambda) = (\lambda^{k+2} h_0(\lambda) e^{-\lambda\phi})^{1/2}$ in the Cauchy equation (37).

To evaluate the $lim_{\phi\to\infty} \frac{m_{k+1}(\phi)}{m_k(\phi)}$, we use a similar argument as for $lim_{\phi\to\infty}\ m_k(\phi)$ in part 2 above, so we include fewer details here. We use the change of variable $z = \lambda\phi$, to express $lim_{\phi\to\infty} \frac{m_{k+1}(\phi)}{m_k(\phi)}$ as

$$\lim{}_{\phi\to\infty} \frac{m_{k+1}(\phi)}{m_k(\phi)} = \lim_{\phi\to\infty} \frac{1}{\phi} \frac{\Gamma(k+2)}{\Gamma(k+1)} = \lim_{\phi\to\infty} \frac{k+1}{\phi}. \tag{38}$$

This implies that $lim_{\phi\to\infty} \frac{m_{k+1}(\phi)}{m_k(\phi)} = 0$ and that $\frac{m_{k+1}(\phi)}{m_k(\phi)} = \frac{k+1}{\phi} + o(\frac{1}{\phi})$ ■.

Proof of Theorem 3.

Part 1. We start with the Bellman equations for the discrete time optimization problem which are equations (17) and (18). From the structure of the problem as shown in (16), the highest possible value of social welfare is $v/r - \underline{c}$ and the lowest possible value is 0. Let S be the space of continuous functions from $\phi \in R_+$ to $[0, v/r - \underline{c}]$. Then S is a complete metric space under the sup norm. We define the Bellman operator $T(\Delta t)$ on functions $W \in S$ as

$$T(\Delta t)W(\phi) = \max_{p\in[\underline{c},\overline{c}]} \left\{ \begin{array}{c} m_1(\phi)\Delta t \int_{\underline{c}}^{p} \left(\frac{v}{r} - c\right) f(c)dc + \\ (1 - m_1(\phi)\Delta t F(p))\,(1 - r\Delta t) W(\phi + \Delta\phi) \end{array} \right\}$$

The fact that W takes values in $[0, v/r - \underline{c}]$ implies that $T(\Delta t)W$ does as well. That $T(\Delta t)W$ is a continuous function of ϕ follows from standard theory given the fact that the feasible set for p is compact.[27] Thus the Bellman operator maps S to S. It is straightforward to check that $T(\Delta t)$ is a contraction mapping, and so has a unique fixed point $V^{\Delta t}(\phi)$ in S. It is also straightforward to check that $T(\Delta t)$ maps decreasing functions to decreasing functions, so that $V^{\Delta t}(\phi)$ is decreasing in ϕ.

[27] See Stokey et al. (1989), Chapter 3.

Subtracting $V(\phi)$ from both sides of (17) and dividing by Δt, we obtain

$$0 = \max_p \left\{ \begin{array}{c} m_1(\phi) \int_{\underline{c}}^p \left(\frac{v}{r} - c\right) f(c)dc + \frac{V(\phi+\Delta\phi)-V(\phi)}{\Delta t} + \\ (-r - m_1(\phi)F(p) + rm_1(\phi)\Delta t F(p))V(\phi + \Delta\phi) \end{array} \right\} \tag{39}$$

$$\text{where } \frac{\Delta\phi}{\Delta t} = F(p) \tag{40}$$

Taking $\Delta t \to 0$, we obtain the Hamilton-Jacobi-Bellman equations[28]

$$0 = \max_p \left\{ m_1(\phi) \int_{\underline{c}}^p \left(\frac{v}{r} - c\right) f(c)dc - (r + m_1(\phi)F(p))\, V(\phi) + \frac{dV}{d\phi}\frac{d\phi}{dt} \right\} \tag{41}$$

$$\text{where } \frac{d\phi}{dt} = F(p) \tag{42}$$

It follows that $\lim_{\Delta t \to 0} V^{\Delta t}(\phi) = V(\phi) \in S$ is the unique solution to (41) and (42). Since $V^{\Delta t}(\phi)$ is decreasing in ϕ, so is $V(\phi)$.

Part 2. In this part, we derive the central equation of motion (24) for the optimal policy and establish that the optimal policy is unique. We also show some properties of continuity, and we show that $\lim_{\phi \to \infty} V(\phi) = 0$.

We first establish that $p(\phi)$ is unique. Substituting $p(\phi)$ into the HJB equation (19) and using integration by parts we have

$$\begin{aligned} 0 \;=\; & m_1(\phi)F(p(\phi))\left(\frac{v}{r} - p(\phi)\right) + m_1(\phi)\int_{\underline{c}}^{p(\phi)} F(c)dc \\ & -(r + m_1(\phi)F(p(\phi)))V(\phi) + F(p(\phi))\frac{dV(\phi)}{d\phi} \end{aligned} \tag{43}$$

Above, we derived a first order condition (21) for $p(\phi)$:

$$V(\phi) \;= \left(\frac{v}{r} - p(\phi)\right) + \frac{1}{m_1(\phi)}\frac{dV(\phi)}{d\phi} \tag{44}$$

Using (44) to eliminate $\frac{d}{d\phi}V(\phi)$, we simplify (43) as:

$$V(\phi) = \frac{m_1(\phi)}{r}\int_{\underline{c}}^{p(\phi)} F(c)dc \tag{45}$$

[28]In this derivation, we use the relationship $\lim_{\Delta t \to 0} \frac{V(\phi+\Delta\phi)-V(\phi)}{\Delta t} = \lim_{\Delta t \to 0} \frac{V(\phi+F(p)\Delta t)-V(\phi)}{\Delta t} = \frac{dV}{d\phi}F(p) = \frac{dV}{d\phi}\frac{d\phi}{dt}$.

Equation (45) is the same as equation (22) above. Because $V(\phi)$ is unique, (45) implies that $p(\phi)$ is unique.

That $\lim_{\phi\to\infty} V(\phi) = 0$ now follows from (45) and the fact that $\lim_{\phi\to\infty} m_1(\phi) = 0$ (Lemma 2).

We next analyze the derivatives of $V(\phi)$ and $p(\phi)$. From (45), the derivative of $V(\phi)$ is

$$\frac{dV(\phi)}{d\phi} = \frac{1}{r}\left\{\frac{dm_1(\phi)}{d\phi}\int_{\underline{c}}^{p(\phi)} F(c)dc + m_1(\phi)F(p(\phi))\frac{dp(\phi)}{d\phi}\right\} \tag{46}$$

Using (45) and (46) to eliminate $V(\phi)$ and $\frac{d}{d\phi}V(\phi)$ in (43), we obtain

$$F(p(\phi))\frac{dp(\phi)}{d\phi} = \frac{1}{m_1(\phi)}(m_1^2(\phi) - \frac{dm_1(\phi)}{d\phi})\int_{\underline{c}}^{p(\phi)} F(c)dc - r(\frac{v}{r} - p\,(\phi)) \tag{47}$$

Using $\frac{dm_1(\phi)}{d\phi} = m_{1.}^2(\phi) - m_2(\phi)$, we write (47) as

$$\frac{dp(\phi)}{d\phi} = \frac{1}{F(p(\phi))}(\frac{m_2(\phi)}{m_1(\phi)}\int_{\underline{c}}^{p(\phi)} F(c)dc - r(\frac{v}{r} - p\,(\phi))) \tag{48}$$

This is the same as equation (24) above and is our central equation of motion. Equation (24) implies that $p(\phi)$ is continuously differentiable. This is clear as long as $F(p(\phi) \neq 0$. However, if $F(p(\phi)) = 0$, then $p(\phi) = \underline{c}$, and (45) implies that $V(p(\phi)) = V(\underline{c}) = 0$. This contradicts the optimality of V because any constant policy p with $\underline{c} < p < \frac{v}{r}$ yields strictly positive value. It follows that $\frac{dp(\phi)}{d\phi}$ is continuous. A similar argument shows that $\frac{d^2p(\phi)}{d\phi^2}$ is continuous. Since $p(\phi)$ is continuously differentiable, equation (46) implies that $V(\phi)$ is as well.

Part 3. In this part, we prove that the optimal reward policy $p(\phi)$ is strictly increasing with $\lim_{\phi\to\infty} p(\phi) = \frac{v}{r}$.

We first observe that $p(\phi) \leq \frac{v}{r}$. This follows from the result in part 1 above that $V(\phi)$ is decreasing in ϕ. For any reward policy $p_1 > \frac{v}{r}$, we compare the welfare that it yields to the welfare under the reward policy $p_2 = \frac{v}{r}$. In the Bellman equation (17), in the current period, social welfare is strictly higher under $p_2 = \frac{v}{r}$, because p_1 allows ideas with cost between $\frac{v}{r}$ and $\bar{c}$ to be invested, and these have negative welfare. In the future period, social welfare is also higher under $p_2 = \frac{v}{r}$, because p_2 generates less learning ($\Delta\phi_2 = F(\frac{v}{r})\Delta t \leq F(p_1)\Delta t = \Delta\phi_1$)

and V is decreasing in ϕ. So $p_2 = \frac{v}{r}$ yields strictly higher welfare than p_1, and p_1 cannot be optimal.

We next prove a lemma about the behavior of p.

Lemma 3 *For any ϕ_1 such that $p'(\phi_1) = 0$, we must have that $p''(\phi_1) < 0$.*

Proof of Lemma 6. We define a function $G(p, \phi)$ as follows:

$$G(p, \phi) = \frac{1}{F(p)}(\frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p} F(c)dc - r(\frac{v}{r} - p))$$

From (48), we have $\frac{dp(\phi)}{d\phi} = G(p(\phi), \phi)$. Differentiating (48), we have:

$$\begin{aligned} p''(\phi) &= G_p(p, \phi)p'(\phi) + G_\phi(p, \phi) \\ &= G_p(p, \phi)p'(\phi) + \frac{1}{F(p)}(\frac{d}{d\phi}\frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p} F(c)dc) \end{aligned}$$

At ϕ_1, since $p'(\phi_1) = 0$, we have

$$p''(\phi_1) = \frac{1}{F(p(\phi_1))}(\frac{d}{d\phi}\frac{m_2(\phi_1)}{m_1(\phi_1)} \int_{\underline{c}}^{p(\phi_1)} F(c)dc)$$

From Lemma 2, we have $\frac{d}{d\phi}\frac{m_2(\phi_1)}{m_1(\phi_1)} < 0$. Since $p(\phi_1) > \underline{c}$ at any optimal policy, this implies that $p''(\phi_1) < 0$. ■

We next prove that if $p'(\phi_1) < 0$ for some ϕ_1, then $p'(\phi) < 0$ for all $\phi > \phi_1$. Suppose this is not true. Then there is some ϕ_2 with $\phi_2 > \phi_1$ and $p'(\phi_2) \geq 0$. Since p and p' are continuously differentiable, there must exist an inflection point ϕ_3 with $\phi_1 < \phi_3 \leq \phi_2$ such that $p'(\phi_3) = 0$ and $p''(\phi_3) \geq 0$. But this contradicts Lemma 6.

If $p'(\phi) < 0$ for all $\phi > \phi_1$, then since $p(\phi)$ lies in the compact interval $[\underline{c}, \frac{v}{r}]$, it must converge downward to a limit $p_{\min} < \frac{v}{r}$ as $\phi \to \infty$. So $\lim_{\phi\to\infty} p(\phi) = p_{\min}$ and $\lim_{\phi\to\infty} \frac{dp(\phi)}{d\phi} = 0$. Using (24) and recalling from Lemma 2 that $\lim_{\phi\to\infty} \frac{m_2(\phi)}{m_1(\phi)} = 0$, we have

$$\begin{aligned} 0 &= \lim_{\phi\to\infty} \frac{dp(\phi)}{d\phi} = \lim_{\phi\to\infty} \frac{1}{F(p(\phi))}\left(\frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p(\phi)} F(c)dc - r(\frac{v}{r} - p(\phi))\right) && (49) \\ &= \frac{1}{F(p_{\min})}\left(-r(\frac{v}{r} - p_{\min})\right) && (50) \end{aligned}$$

If $p_{\min} = \underline{c}$, the right-hand side of (50) diverges, so this does not solve the equation. If $p_{\min} > \underline{c}$, then since $p_{\min} < \frac{v}{r}$, the right-hand side of (50) is strictly negative, so the limit again cannot be 0. These contradictions imply that there cannot be a ϕ_1 such that $p'(\phi_1) < 0$. That is,

$$p'(\phi) \geq 0 \text{ for all } \phi$$

Since $p(\phi)$ lies in the compact interval $[\underline{c}, \frac{v}{r}]$, $p(\phi)$ must converge upward to a limit $p^{\max} \leq \frac{v}{r}$ as $\phi \to \infty$. So we have $\lim_{\phi\to\infty} p(\phi) = p^{\max}$ and $\lim_{\phi\to\infty} \frac{dp(\phi)}{d\phi} = 0$. From equation (49), $\lim_{\phi\to\infty} p'(\phi) = \frac{1}{F(p^{\max})}(-r(\frac{v}{r} - p^{\max}) = 0$, and this can only hold if $p^{\max} = v/r$.■

Part 4. We have shown that $p(\phi)$ is an increasing function, with $\lim_{\phi\to\infty} p(\phi) = v/r$. In this part, show that $p(\phi) = \frac{v}{r} - \frac{2}{\phi r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc + o(\frac{1}{\phi})$.

Rearranging the equation of motion (24), we have

$$\frac{F(p(\phi))}{r}\left(\phi\frac{dp(\phi)}{d\phi}\right) = \frac{\phi}{r}\frac{m_2(\phi)}{m_1(\phi)}\int_{\underline{c}}^{p(\phi)} F(c)dc - \phi(\frac{v}{r} - p(\phi)) \tag{51}$$

Using $\lim_{\phi\to\infty} p(\phi) = \frac{v}{r}$ and $\frac{m_2(\phi)}{m_1(\phi)} = \frac{2}{\phi} + o(\frac{1}{\phi})$ (Lemma 2), we obtain

$$\frac{F(\frac{v}{r})}{r}\lim_{\phi\to\infty}\left(\phi\frac{dp(\phi)}{d\phi}\right) = \frac{2}{r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc - \lim_{\phi\to\infty}\Big(\phi(\frac{v}{r} - p(\phi))\Big) \tag{52}$$

Using l'Hôpital's rule, we have

$$\lim_{\phi\to\infty}\Big(\phi(\frac{v}{r} - p(\phi))\Big) = \lim_{\phi\to\infty}\left(\frac{\frac{v}{r} - p(\phi)}{\frac{1}{\phi}}\right) = \lim_{\phi\to\infty}\left(\phi^2\frac{dp(\phi)}{d\phi}\right) \tag{53}$$

Substituting (53) into (52), and rearranging terms, we have

$$\frac{2}{r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc = \lim_{\phi\to\infty}\left(\frac{F(\frac{v}{r})}{r}\phi\frac{dp(\phi)}{d\phi} + \phi^2\frac{dp(\phi)}{d\phi}\right) = \lim_{\phi\to\infty}\left(\phi^2\frac{dp(\phi)}{d\phi}(1 + \frac{F(\frac{v}{r})}{r\phi})\right)$$

Since $\lim_{\phi\to\infty}(1 + \frac{F(\frac{v}{r})}{r\phi}) = 1$, we now have

$$\frac{2}{r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc = \lim_{\phi\to\infty}\left(\phi^2\frac{dp(\phi)}{d\phi}\right) = \lim_{\phi\to\infty}\Big(\phi(\frac{v}{r} - p(\phi))\Big) \tag{54}$$

where the last equality comes from (53). Using the definition of $o(\cdot)$, this establishes that

$$p(\phi) = \frac{v}{r} - \frac{2}{\phi r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc + o(\frac{1}{\phi}) \tag{55}$$

$\blacksquare$

Proof of Theorem 4.

Part 1. We use $p_{na}(\phi)$ to denote the optimal policy in the static model when the arrival rate is known to be $m_1(\phi)$. From equation (8), $p_{na}(\phi)$ is the unique solution to

$$\frac{v}{r} - p_{na}(\phi) = \frac{m_1(\phi)}{r} \int_{\underline{c}}^{p_{na}(\phi)} F(c)dc. \tag{56}$$

From the implicit function theorem, $p_{na}(\phi)$ is a continuously differentiable function of ϕ. From Theorem 3, $p(\phi)$ is also continuously differentiable. To compare the policies, we first prove a single crossing property.

Lemma 4 *[Single-crossing property] If $p(\phi_c) = p_{na}(\phi_c)$ for some $\phi_c \geq 0$, then $p'(\phi_c) > p'_{na}(\phi_c)$.*

Proof of Lemma 4. Let ϕ_c be a crossing point, so that $p(\phi_c) = p_{na}(\phi_c)$. We first derive an expression for the derivative of $p_{na}(\phi)$. Differentiating both sides of (56) with respect to ϕ we have

$$-\frac{dp_{na}(\phi)}{d\phi} = \frac{1}{r}\frac{dm_1(\phi)}{d\phi} \int_{\underline{c}}^{p_{na}(\phi)} F(c)dc + \frac{m_1(\phi)}{r} F(p_{na}(\phi)) \frac{dp_{na}(\phi)}{d\phi}$$

Rearranging we have

$$\frac{dp_{na}(\phi)}{d\phi} = -\frac{\frac{dm_1(\phi)}{d\phi}\int_{\underline{c}}^{p_{na}(\phi)} F(c)dc}{(r + m_1(\phi)F(p_{na}(\phi))} \tag{57}$$

Substituting $\frac{dm_1(\phi)}{d\phi} = m_1^2(\phi) - m_2(\phi)$ into (57) we have

$$\frac{dp_{na}(\phi)}{d\phi} = \frac{\left(m_2(\phi) - m_1^2(\phi)\right)}{(r + m_1(\phi)F(p_{na}(\phi))} \int_{\underline{c}}^{p_{na}(\phi)} F(c)dc \tag{58}$$

We next derive an expression for the derivative of $p(\phi)$. From (24), we have

$$\frac{dp(\phi)}{d\phi} = \frac{1}{F(p(\phi))}\left(\frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p(\phi)} F(c)dc - r\left(\frac{v}{r} - p\left(\phi\right)\right)\right) \tag{59}$$

Using $p(\phi_c) = p_{na}(\phi_c)$ in equation (56), we obtain

$$p(\phi_c) = \frac{v}{r} - \frac{m_1(\phi_c)}{r} \int_{\underline{c}}^{p(\phi_c)} F(c)dc \tag{60}$$

Finally, substituting this expression for $p(\phi_c)$ into the last term in (59), we have

$$\begin{aligned}
\frac{dp(\phi)}{d\phi}|_{\phi_c} &= \frac{1}{F(p(\phi_c))}(\frac{m_2(\phi_c)}{m_1(\phi_c)}\int_{\underline{c}}^{p(\phi_c)} F(c)dc - r(\frac{v}{r}) + r(\frac{v}{r} - \frac{m_1(\phi_c)}{r}\int_{\underline{c}}^{p(\phi_c)} F(c)dc)) \\
&= \frac{\left(m_2(\phi_c) - m_1^2(\phi_c)\right)}{m_1(\phi_c)F(p(\phi_c))}\int_{\underline{c}}^{p(\phi_c)} F(c)dc > \frac{dp_{na}(\phi)}{d\phi}|_{\phi_c} \qquad (61)
\end{aligned}$$

where the inequality in (61) follows from (58).■

Part 2 In this part, we show that $p_{na}(\phi) > p(\phi)$ for all ϕ.

We first show that $p_{na}(\phi) > p(\phi)$ when ϕ is large. Since $\lim_{\phi\to\infty} m_1(\phi) = 0$, equation (56) implies that $\lim_{\phi\to\infty} p_{na}(\phi) = \frac{v}{r}$. From Lemma 2, we have $m_1(\phi) = \frac{1}{\phi} + o(\frac{1}{\phi})$ as $\phi \to \infty$. Taking these limits in (56), we obtain

$$p_{na}(\phi) = \frac{v}{r} - \frac{1}{\phi r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc + o(\frac{1}{\phi}) \qquad (62)$$

Along with (55), this implies that

$$p_{na}(\phi) - p\left(\phi\right) = \frac{1}{\phi r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc + o(\frac{1}{\phi})$$

so that

$$\lim_{\phi\to\infty} \phi\left(p_{na}(\phi) - p\left(\phi\right)\right) = \frac{1}{r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc > 0$$

This implies there must be a ϕ_1 such that $\phi\left(p_{na}(\phi) - p\left(\phi\right)\right) > 0$ for all $\phi > \phi_1$, and hence that $p_{na}(\phi) > p\left(\phi\right)$ for all $\phi > \phi_1$.

Finally, we show that $p_{na}(\phi) > p(\phi)$ for all $\phi > 0$. Suppose that this is not true. Then $p_{na}(\phi_2) \leq p(\phi_2)$ for some ϕ_2. Since $p_{na}(\phi) > p\left(\phi\right)$ when ϕ is large, there is some ϕ_c that is the largest ϕ with $p_{na}(\phi) \leq p(\phi)$. By continuity, we must have that $p_{na}(\phi_c) = p\left(\phi_c\right)$. From Lemma 4, $p'_{na}(\phi_c) < p'(\phi_c)$. This is a contradiction because it is inconsistent with $p_{na}(\phi) > p(\phi)$ for $\phi > \phi_c$.■

Proof of Theorem 5:

Part 1. We first derive the social planner's expected value of the arrival rate in the moment after innovation occurs. At ϕ, the social planner's expected arrival rate is $m_1(\phi)$. Using a discrete time approximation and equation (18), we have $\Delta\phi = F(p)\Delta t$. In the interval $\Delta\phi$, one of two events will happen. If innovation does not occur, the state is

$\phi + \Delta\phi$, and the social planner's expected arrival rate is $m_1(\phi + \Delta\phi)$. If innovation occurs, the social planner observes the arrival and updates its beliefs (even though there are no more actions to take). We use $q(\phi + \Delta\phi)$ to denote this expectation. Using the martingale property of Bayesian updating, $q(\phi + \Delta\phi)$ must satisfy

$$m_1(\phi) = (1 - m_1(\phi)\Delta\phi)m_1(\phi + \Delta\phi) + q(\phi + \Delta\phi)m_1(\phi)\Delta\phi \tag{63}$$

Substituting $m_1(\phi + \Delta\phi) = m_1(\phi) + m_1'(\phi)\Delta\phi$ into equation (119) and simplifying, we have

$$0 = m_1'(\phi)\Delta\phi - m_1^2(\phi)\Delta\phi - m_1(\phi)m_1'(\phi)(\Delta\phi)^2 + q(\phi + \Delta\phi)m_1(\phi)\Delta\phi \tag{64}$$

Dividing by $\Delta\phi$ and taking the limit as $\Delta\phi \to 0$, we have

$$m_1'(\phi) - m_1^2(\phi) + q(\phi)m_1(\phi) = 0 \tag{65}$$

From Lemma 2, we have $m_1'(\phi) = (m_1^2(\phi) - m_2(\phi))$. Substituting this into (65), we have

$$q(\phi) = \frac{m_2(\phi)}{m_1(\phi)} \tag{66}$$

This establishes that the social planner's expected arrival rate in the moment after an innovation at ϕ is $q(\phi) = \frac{m_2(\phi)}{m_1(\phi)}$.

Part 2. We use $p^*(\lambda)$ to denote the optimal policy in the static model when the arrival rate is known to be λ. From equation (8), we have that $p^*(\frac{m_2(\phi)}{m_1(\phi)})$ is the unique p^* that solves

$$\frac{v}{r} - p^* = \frac{1}{r}\frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p^*} F(c)dc. \tag{67}$$

From equation (24), the optimal policy $p(\phi)$ in our main model satisfies

$$F(p)\frac{dp}{d\phi} = \frac{m_2(\phi)}{m_1(\phi)} \int_{\underline{c}}^{p} F(c)dc - r(\frac{v}{r} - p) \tag{68}$$

From (67), the right hand side of (68) is 0 at $p = \frac{m_2(\phi)}{m_1(\phi)}$. The right hand side of (68) is stricting increasing in p, so it is positive for $p > p^*(\frac{m_2(\phi)}{m_1(\phi)})$ and negative for $p < p^*(\frac{m_2(\phi)}{m_1(\phi)})$. Since the left hand side of (68) is positive, we must have $p(\phi) > p^*(\frac{m_2(\phi)}{m_1(\phi)})$.■

Proof of Theorem 6: We first develop some properties for the relationship between the cost distributions F and G.

The strict MLRP implies that for any c_1 and c_2 in $[\underline{c}, \overline{c}]$, with $c_2 > c_1$,

$$\frac{f(c_2)}{f(c_1)} > \frac{g(c_2)}{g(c_1)} \tag{69}$$

Integrating from $c_1 = \underline{c}$ to $c_1 = c_2$, we obtain

$$\frac{f(c_2)}{F(c_2)} > \frac{g(c_2)}{G(c_2)} \tag{70}$$

for any $c_2 \in (\underline{c}, \overline{c}]$.

We next define

$$\Phi_F(p) \equiv \frac{\int_{\underline{c}}^{p} F(c)dc}{F(p)} \ \ \textit{and} \ \ \Phi_G(p) = \frac{\int_{\underline{c}}^{p} G(c)dc}{G(p)}. \tag{71}$$

We want to show that $\Phi_F(p) < \Phi_G(p)$ for $p \in (\underline{c}, \overline{c}]$. For any $p \in (\underline{c}, \overline{c}]$, consider some $z \in (\underline{c}, p)$. We have

$$\int_z^p \frac{f(x)}{F(x)} dx = \ln F(p) - \ln F(z)$$

so that $\exp(\int_z^p \frac{f(x)}{F(x)} dx) = \frac{F(p)}{F(z)}$ and

$$\exp(-\int_z^p \frac{f(x)}{F(x)} dx) = \frac{F(z)}{F(p)}. \tag{72}$$

The same argument shows that

$$\exp(-\int_z^p \frac{g(x)}{G(x)} dx) = \frac{G(z)}{G(p)}. \tag{73}$$

From (70), we have

$$\int_z^p \frac{f(x)}{F(x)} dx > \int_z^p \frac{g(x)}{G(x)} dx \tag{74}$$

Combining (72), (73), and (74), we obtain

$$\frac{F(z)}{F(p)} = \exp(-\int_z^p \frac{f(x)}{F(x)} dx) < \exp(-\int_z^p \frac{g(x)}{G(x)} dx) = \frac{G(z)}{G(p)} \tag{75}$$

Integrating the expressions in (75) over $z \in [\underline{c}, p]$ now gives

$$\Phi_F(p) \equiv \frac{\int_{\underline{c}}^{p} F(c)dc}{F(p)} < \frac{\int_{\underline{c}}^{p} G(c)dc}{G(p)} = \Phi_G(p) \text{ for } p \in (\underline{c}, \overline{c}]. \tag{76}$$

which is what we wanted to show.

We next establish a single crossing property for the optimal policies $p_F(\phi)$ and $p_G(\phi)$ under F and G.

Lemma 5 *[Single crossing property] If* $p_F(\phi_c) = p_G(\phi_c)$*, then* $p'_F(\phi_c) < p'_G(\phi_c)$.

Proof of Lemma 5. From Theorem 3, $p_F(\phi)$ and $p_G(\phi)$ satisfy the equation (24):

$$\frac{dp_I(\phi)}{d\phi} = \frac{1}{I(p_I(\phi))}(\frac{m_2(\phi)}{m_1(\phi)}\int_{\underline{c}}^{p_I(\phi)} I(c)dc - r(\frac{v}{r} - p_I(\phi)))for\ I \in \{F, G\}$$

Using Φ_F and Φ_G as defined in (71), we can write this as

$$\frac{dp_I(\phi)}{d\phi} = \frac{m_2(\phi)}{m_1(\phi)}\Phi_I(p_I(\phi)) - \frac{(v - rp_I(\phi))}{F(p_I(\phi))}for\ I \in \{F, G\}$$

Using (76), we have

$$\begin{aligned} \frac{dp_F(\phi)}{d\phi} &= \frac{m_2(\phi)}{m_1(\phi)}\Phi_F(p_F(\phi)) - \frac{(v - rp_F(\phi))}{F(p_F(\phi))} \\ &< \frac{m_2(\phi)}{m_1(\phi)}\Phi_G(p_F(\phi)) - \frac{(v - rp_F(\phi))}{F(p_F(\phi))} \end{aligned} \tag{77}$$

Let ϕ_c be a crossing point, so that $p_F(\phi_c) = p_G(\phi_c)$. Substituting ϕ_c into (77) we have

$$\frac{dp_F(\phi)}{d\phi}|_{\phi_c} < \frac{m_2(\phi_c)}{m_1(\phi_c)}\Phi_G(p_G(\phi_0)) - \frac{(v - rp_G(\phi_c))}{F(p_G(\phi_c))} = \frac{dp_G(\phi)}{d\phi}|_{\phi_c}$$

which completes the proof.■

We next show that $p_F(\phi) > p_G(\phi)$ for ϕ sufficiently large. In Theorem 3, we derived the asymptotic behavior of the optimal policy. Applying this separately to p_F and p_G, we have

$$\begin{aligned} p_F(\phi) &= \frac{v}{r} - \frac{2}{\phi r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc + o(\frac{1}{\phi}) \\ p_G(\phi) &= \frac{v}{r} - \frac{2}{\phi r}\int_{\underline{c}}^{\frac{v}{r}} G(c)dc + o(\frac{1}{\phi}) \end{aligned}$$

These equations imply that

$$\lim_{\phi\to\infty} \phi(p_G(\phi) - p_F(\phi)) = \frac{2}{r}\int_{\underline{c}}^{\frac{v}{r}} (F(c) - G(c))dc < 0 \tag{78}$$

where the inequality in (78) follows from the well-known property that the MLRP implies first order stochastic dominance. Specifically, the strict MLRP implies that F dominates G in the sense of strict first order stochastic dominance on $(\underline{c}, \overline{c})$. The inequality (78) implies that there exists some ϕ_1 such that $p_G(\phi) < p_F(\phi)$ for $\phi > \phi_1$.

We can now show that $p_F(\phi) > p_G(\phi)$ for all ϕ. Suppose this is not true. Then $p_F(\phi_2) \leq p_G(\phi_2)$ for some ϕ_2. Since $p_F(\phi) > p_G(\phi)$ for $\phi > \phi_1$, there is some ϕ_c that is the largest value of ϕ for which $p_F(\phi) \leq p_G(\phi)$. From continuity, we must have $p_F(\phi_c) = p_G(\phi_c)$. From Lemma 5, we have $p'_F(\phi_c) < p'_G(\phi_c)$. This is a contradiction, because it is inconsistent with $p_F(\phi) > p_G(\phi)$ for $\phi > \phi_c$. It follows that $p_G(\phi) < p_F(\phi)$ for all ϕ.

Finally, we show that social welfare (the value function) is higher under the cost distribution G than under F. From equation (22), applied to F and G, we obtain

$$V_G(\phi) - V_F(\phi) = \frac{m_1(\phi)}{r}\int_{\underline{c}}^{\frac{v}{r}} G(c)dc - \frac{m_1(\phi)}{r}\int_{\underline{c}}^{\frac{v}{r}} F(c)dc \tag{79}$$

$$= \frac{m_1(\phi)}{r}\int_{\underline{c}}^{\frac{v}{r}} (G(c)-F(c))dc - \frac{m_1(\phi)}{r}\int_{p_G(\phi)}^{\frac{v}{r}} G(c)dc + \frac{m_1(\phi)}{r}\int_{p_F(\phi)}^{\frac{v}{r}} F(c)dc \tag{80}$$

The second and third terms on the right-hand side of (80) are $O(\frac{1}{\phi})$ as $\phi \to \infty$. To see this for the second term, we use l'Hôpital's rule, rule to compute

$$\begin{aligned}\lim_{\phi\to\infty} \phi \int_{p_G(\phi)}^{\frac{v}{r}} G(x)dx &= \lim_{\phi\to\infty} \frac{\int_{p_G(\phi)}^{\frac{v}{r}} G(x)}{\frac{1}{\phi}} = \lim_{\phi\to\infty} \frac{-G(p_G(\phi))\frac{dp_G(\phi)}{d\phi}}{\frac{-1}{\phi^2}} \\ &= G(\frac{v}{r})\lim_{\phi\to\infty}\phi^2\frac{dp_G(\phi)}{d\phi} = G(\frac{v}{r})\frac{2}{r}\int_{\underline{c}}^{\frac{v}{r}} G(c)dc \end{aligned}\tag{81}$$

where the last equality in (81) comes from equation (54). Since $G(\frac{v}{r})\frac{2}{r}\int_{\underline{c}}^{\frac{v}{r}} G(c)dc \neq 0$, we have $\int_{p_G(\phi)}^{\frac{v}{r}} G(x)dx = O(\frac{1}{\phi})$. The same argument shows that $\int_{p_F(\phi)}^{\frac{v}{r}} F(x)dx = O(\frac{1}{\phi})$. From Lemma 2, we have $m_1(\phi) = \frac{1}{\phi} + o(\frac{1}{\phi})$. Using these asymptotics in equation (79) with the rules of operating with $o(\frac{1}{\phi})$ and $O(\frac{1}{\phi})$, we have that as $\phi \to \infty$

$$V_G(\phi) - V_F(\phi) = \frac{1}{r}(\frac{1}{\phi} + o(\frac{1}{\phi}))\int_{\underline{c}}^{\frac{v}{r}} (G(c)-F(c))dc + o(\frac{1}{\phi}) \tag{82}$$

$$= \frac{1}{r\phi}\int_{\underline{c}}^{\frac{v}{r}} (G(c)-F(c))dc + o(\frac{1}{\phi}) \tag{83}$$

Multiplying by $r\phi$ and taking the limit, we now have

$$\lim_{\phi\to\infty} r\phi(V_G(\phi) - V_F(\phi)) = \int_{\underline{c}}^{\frac{v}{r}} (G(c)-F(c))dc > 0. \tag{84}$$

where the last inequality comes from the fact that F stochastically dominates G. Therefore, there is some ϕ_1 such that $V_G(\phi) > V_F(\phi)$ for $\phi > \phi_1$.

It remains to show that $V_G(\phi) > V_F(\phi)$ for all ϕ. Suppose to the contrary that there exists a ϕ_2 such that $V_G(\phi_2) \leq V_F(\phi_2)$. From the social planner's first order condition (21), we have

$$
\begin{aligned}
m_1(\phi)(V_F(\phi) + \ p_F(\phi) - \frac{v}{r}) &= \frac{dV_F(\phi)}{d\phi} && (85)\\
m_1(\phi)(V_G(\phi) + \ p_G(\phi) - \frac{v}{r}) &= \frac{dV_G(\phi)}{d\phi} && (86)
\end{aligned}
$$

Since $V_G(\phi) > V_F(\phi)$ for large ϕ, there must be some largest value of ϕ, which we will call ϕ_c, such that V_G and V_F have the same value. For $\phi > \phi_c$, we then have $V_G(\phi) > V_F(\phi)$. We have already proved that $p_F(\phi) > p_G(\phi)$ for all ϕ, and so in particular $p_F(\phi_c) > p_G(\phi_c)$. Therefore (85) and (86) imply that $V_F'(\phi_c) > V_G'(\phi_c)$. Since V_F' and V_G' are negative, this means that V_F crosses V_G from below at ϕ_c which is inconsistent with $V_G(\phi) > V_F(\phi)$ for all $\phi > \phi_c$. This contradiction establishes that $V_G(\phi) > V_F(\phi)$ for all ϕ. ∎

ONLINE APPENDIX

In this appendix, we assume that the social planner observes the arrival of ideas before deciding whether to fund the investment. Otherwise, the model is unchanged.

We first discuss the state space and the social planner's beliefs about the arrival rate λ. We then analyze the dyanmic optimization problem that the social planner faces.

1 State space and beliefs

The arrival of ideas contains information that informs the the social planner's beliefs about the arrival rate λ. The social planner's initial beliefs are given by the probability distribution $h_0(\lambda)$. Intuitively, the more ideas that arrive between time 0 and time t, the more optimistic the planner is about the arrival rate. Under the Poisson arrival process, the timing of the arrivals of ideas within the interval $[0, t]$ does not matter for statistical inference.[29] Thus, a sufficient statistic for the social planner's beliefs is (k, t) where k is the number of ideas that have arrived in the interval $[0, t]$. We use $h(\cdot, k, t)$ to denote the social planner's beliefs at (k, t) where $h(\lambda, 0, 0) = h_0(\lambda)$. The expected value of λ is $m_1(k, t) = \int_0^\infty \lambda h(\lambda, k, t) d\lambda$.

At (k, t), the social planner's beliefs are related to the initial beliefs $h_0(\lambda)$ as follows. For a given value of λ, as seen from time $t = 0$, the Poisson probability that exactly k ideas arrive by time t is[30]

$$q_\lambda(k, t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \tag{87}$$

Applying Bayes' Rule, the social planner's belief at (k, t) are

$$h(\lambda, k, t) = \frac{q_\lambda(k, t) h_0(\lambda)}{\int_0^\infty q_{\lambda'}(k, t) h_0(\lambda') d\lambda'} \tag{88}$$

The n-th moment of $h(\lambda, k, t)$ is

$$m_n(k, t) = \int_0^\infty \lambda^n h(\lambda, k, t) d\lambda \tag{89}$$

and the cumulative distribution is $H(\lambda, k, t) = \int_0^\lambda \lambda' h(\lambda', k, t) d\lambda'$. The social planner may become either more or less optimistic over time. The next lemma shows that the social planner becomes more pessimistic about the arrival rate when ideas do not arrive and more optimistic when one does.

Lemma 6 *The cumulative distribution $H(\lambda, k, t)$ is decreasing in t and increasing in k in the sense of first order stochastic dominance.*

[29]The costs of the ideas do not matter for statistical inference because the arrival rate is independent of the cost distributon $F(c)$.

[30]Note that $\sum_{k=0}^\infty q_\lambda(k, t) = 1$.

2 Optimal policy

We now formulate the social planner's optimization problem. At time t if an idea arrives, the the social planner observes the cost c. The social planner updates its beliefs to reflect the arrival, and then decides whether to invest the idea or let it go. We use (k,t,c) to denote this state, where k is the number of ideas that have arrived including the one that just arrived. If no idea arrives, or if the social planner has let an idea go, we denote the state by $(k,t,\emptyset)$.

To solve the model, we apply standard techniques from dyanmic programming. We approximate continuous time with a discrete time interval Δt, solve that problem and then take the limit of the solution (social welfare) as $\Delta t \to 0$. The discount factor e^{-rt} is approximated[31] by $(1-r\Delta t)$. There are two Bellman equations corresponding to the states (k,t,c) and $(k,t,\emptyset)$ respectively. At (k,t,c), the Bellman equation is

$$V(k,t,c) = \max\{\frac{v}{r} - c, V(k,t,\emptyset)\} \tag{90}$$

where $V(k,t,\emptyset)$ is the continuation value if the social planner does not invest the idea.

At $(k,t,\emptyset)$, the Bellman equation is

$$\begin{aligned} V(k,t,\emptyset) = (1-r\Delta t)\{m_1(k,t)\Delta t \int_{\underline{c}}^{\overline{c}} V(k+1,t+\Delta t,c)f(c)dc \\ + (1-m_1(k,t)\Delta t)V(k,t+\Delta t,\emptyset)\} \end{aligned} \tag{91}$$

Equation (91) is the social planner's continuation value when there is no idea to invest. The first term on the right hand side of (91) corresponds to the event that an idea arrives in the next period, $t+\Delta t$, which happens with probability $m_1(k,t)\Delta t$. The second term corresponds to the event that no idea arrives at $t+\Delta t$ which happens with probability $(1-m_1(k,t)\Delta t)$. Both terms are discounted by $(1-r\Delta t)$ as the payoffs do not arise until the next period.

From equation (90), the social planner invests an idea with cost c if and only if

$$\frac{v}{r} - c \geq V(k,t,\emptyset) \tag{92}$$

The social planner can implement this solution by offering a reward policy $p(k,t)$ such at

$$\frac{v}{r} - p(k,t) = V(k,t,\emptyset) \tag{93}$$

The reward policy does not depend on the cost c because the continuation value of not investing does not depend on c.

In the proof of Theorem 7, we derive the Hamilton-Jacobi -Bellman equations

$$V(k,t,c) = \max\{\frac{v}{r} - c, V(k,t,\emptyset)\} \tag{94}$$

[31]Using the Taylor series expansion, we have $e^{-rt} = (1-r\Delta t) + o(\Delta t)$.

and

$$0 = m_1(k,t)\Big\{\int_{\underline{c}}^{p(k+1,t)}(\frac{v}{r}-c)f(c)dc + (1-F(p(k+1,t)))V(k+1,t,\emptyset)\Big\} + \frac{\partial V(k,t,\emptyset)}{\partial t} - V(k,t,\emptyset)(m_1(k,t)+r) \tag{95}$$

The first equation (94) is the same as (90). The second equation (95) is derived by taking the limit of (91) as $\Delta t \to 0$. Using the relationship between the value function and the optimal policy in equation (93), we also derive an equation of motion for the optimal policy

$$\frac{\partial p(k,t)}{\partial t} = m_1(k,t)\int_{\underline{c}}^{p(k+1,t)} F(c)dc + m_1(k,t)(p(k,t)-p(k+1,t)) - r(\frac{v}{r}-p(k,t)) \tag{96}$$

Theorem 7 *The social welfare function, $V(k,t,c)$ and $V(k,t,\emptyset)$, is the unique solution to the social planner's optimization problem which is given by equations (94) and (95). Social welfare is increasing in k and decreasing in t and c. The optimal policy $p(k,t)$ is the unique solution to equation (93). The optimal policy is decreasing in k, increasing in t, and satisfies the equation of motion (96).*

Finally, we compare the optimal policy $p(k,t)$ to the policy that a social planner would choose if the arrival rate λ were fixed at the expected value $m_1(k,t)$. From Theorem 1 in the paper, the optimal policy when λ is known p^* is decreasing in λ. We write $p^*(\lambda)$ to track how p^* depends on the arrival rate. In general, the relationship between $p(k,t)$ and $p^*(m_1(k,t))$ is ambiguous. However, if $p^*(\lambda)$ is convex, we have the following result.

Theorem 8 *If $p^*(\lambda)$ is convex in λ, the optimal reward policy $p(k,t)$ is higher than it would be if the arrival rate were fixed at $\lambda = m_1(k,t)$.*

3 Proofs

Proof of Lemma 6:

Part 1. In this part, we show that the cumulative distribution $H(\lambda,k,t)$ is decreasing in t in the sense of first-order stochastic dominance.

We want to show that $\frac{d}{dt}H(\lambda,k,t) > 0$. Substituting (87) into (88), we have

$$h(\lambda,k,t) = \frac{\lambda^k e^{-\lambda t}h_0(\lambda)}{\int_0^\infty (\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'} \tag{97}$$

Taking the derivative, we have

$$\begin{aligned}
\frac{dh(\lambda,k,t)}{dt} &= \frac{-\lambda(\lambda^k)e^{-\lambda t}h_0(\lambda)}{\int_0^\infty(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'} - h(\lambda,k,t)\frac{(-\int_0^\infty \lambda'(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda')}{\int_0^\infty(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'} \qquad (98)\\
&= -\lambda h(\lambda,k,t) + h(\lambda,k,t)\frac{(\int_0^\infty \lambda'(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda')}{\int_0^\infty(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'}\\
&= h(\lambda,k,t)(-\lambda + \frac{t^k(\int_0^\infty \lambda'(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda')}{t^k\int_0^\infty(\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'})\\
&= h(\lambda,k,t)(-\lambda + \frac{\int_0^\infty \lambda'(t\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'}{\int_0^\infty(t\lambda')^k e^{-\lambda' t}h_0(\lambda')d\lambda'})\\
&= h(\lambda,k,t)(-\lambda + \frac{\int_0^\infty \lambda' q_{\lambda'}(k,t)h_0(\lambda')d\lambda'}{\int_0^\infty q_{\lambda'}(k,t)h_0(\lambda')d\lambda'})\\
&= h(\lambda,k,t)(-\lambda + m_1(k,t)) \qquad (99)
\end{aligned}$$

where $q_{\lambda'}(k,t)$ is given by (87), $h(\lambda,k,t)$ is given by (88), and $m_1(k,t)$ is the expected value of $h(\lambda,k,t)$. This implies that $\frac{dh(\lambda,k,t)}{dt}$ is positive for $\lambda < m_1(k,t)$ and negative for $\lambda > m_1(k,t)$.[32]

From here, we can show that $\frac{d}{dt}H(\lambda,k,t) > 0$ as follows. For $\lambda < m_1(k,t)$, we have

$$\frac{d}{dt}H(\lambda,k,t) = \frac{d}{dt}\int_0^\lambda h(\lambda',k,t)d\lambda' = \int_0^\lambda \frac{d}{dt}h(\lambda',k,t)d\lambda' > 0$$

since $\frac{d}{dt}h(\lambda',k,t) > 0$ for $\lambda' < m_1(k,t)$.

For $\lambda > m_1(k,t)$,we have

$$\begin{aligned}
\frac{d}{dt}H(\lambda,k,t) &= \frac{d}{dt}\int_0^\lambda h(\lambda',k,t)d\lambda' = \frac{d}{dt}\int_0^\infty h(\lambda',k,t)d\lambda' - \frac{d}{dt}\int_\lambda^\infty h(\lambda',k,t)d\lambda'\\
&= 0 - \int_\lambda^\infty \frac{d}{dt}h(\lambda',k,t)d\lambda' > 0
\end{aligned}$$

since $\frac{d}{dt}h(\lambda',k,t) < 0$ for $\lambda' > m_1(k,t)$. In the last line we used the fact that $\int_0^\infty h(\lambda,k,t)d\lambda = 1$ implies $\frac{d}{dt}\int_0^\infty h(\lambda,k,t)d\lambda = 0$.

Therefore, $\frac{d}{dt}H(\lambda,k,t) > 0$ for all (λ,k,t).

Part 2. In this part, we show that the cumulative distribution $H(\lambda,k,t)$ is increasing in k in the sense of first-order stochastic dominance. It follows that the moments $m_n(k,t)$ are increasing in k.

[32] The expected value m_1 is positive since the support of λ is $[0,\infty]$. The only way we could have $m_1 = 0$, is if the social planner is certain that the arrival rate is 0, so that $h(\lambda,\phi)$ has a point mass at $\lambda = 0$. In this degnerate scenario, the beliefs remain constant over time and the optimal policy is indeterminant since there is no chance that the market niche will be filled.

Although k is a discrete variable, we treat it as continuous for the purpose of this proof. Analogously to Part 1, we will show that there is a value λ'such that $\frac{dh(\lambda,k,t)}{dk}$ is negative for $\lambda < \lambda'$and positive for $\lambda > \lambda'$. Recall that

$$h(\lambda, k, t) = \frac{\lambda^k e^{-\lambda t} h_0(\lambda)}{\int_0^\infty (\lambda')^k e^{-\lambda' t} h_0(\lambda') d\lambda'} \tag{100}$$

We first take the derivative of the denominator and numerator of $h(\lambda, k, t)$ separately.

$$\begin{aligned} \frac{d}{dk} \lambda^k e^{-\lambda t} h_0(\lambda) &= \lambda^k \ln(\lambda) e^{-\lambda t} h_0(\lambda) \\ \frac{d}{dk} \int_0^\infty (\lambda')^k e^{-\lambda' t} h_0(\lambda') d\lambda' &= \int_0^\infty (\lambda')^k \ln(\lambda') e^{-\lambda' t} h_0(\lambda') d\lambda' \end{aligned}$$

Taking the derivative of $h(\lambda, k, t)$ using the quotient rule, we have

$$\begin{aligned} \frac{dh(\lambda,k,t)}{dk} &= \frac{\lambda^k \ln(\lambda) e^{-\lambda t} h_0(\lambda)}{\int_0^\infty (\lambda')^k e^{-\lambda' t} h_0(\lambda') d\lambda'} - h(\lambda,k,t) \frac{\int_0^\infty (\lambda')^k \ln(\lambda') e^{-\lambda' t} h_0(\lambda') d\lambda'}{\int_0^\infty (\lambda')^k e^{-\lambda' t} h_0(\lambda') d\lambda'} \\ &= \ln(\lambda) h(\lambda,k,t) - h(\lambda,k,t) \frac{\int_0^\infty (\lambda')^k \ln(\lambda') e^{-\lambda' t} h_0(\lambda') d\lambda'}{\int_0^\infty (\lambda')^k e^{-\lambda' t} h_0(\lambda') d\lambda'} \\ &= h(\lambda,k,t) \left(\ln(\lambda) - \frac{\int_0^\infty (\lambda')^k \ln(\lambda') e^{-\lambda' t} h_0(\lambda') d\lambda'}{\int_0^\infty (\lambda')^k e^{-\lambda' t} h_0(\lambda') d\lambda'} \right) \end{aligned} \tag{101}$$

The right hand side of (101) is negative at $\lambda = 0$ because $\ln(0) = -\infty$ and the second term inside the parentheses is finite. The full expression inside the parentheses is increasing in λ, because $\ln(\lambda)$ is increasing and the second term is constant. Because λ increases to ∞, there exists some $\lambda' \in (0, \infty)$ such that $\frac{dh(\lambda,k,t)}{dk} < 0$ for $\lambda < \lambda'$ and $\frac{dh(\lambda,k,t)}{dk} > 0$ for $\lambda > \lambda'$.

We now show that $\frac{d}{dk} H(\lambda, k, t) < 0$. For $\lambda < \lambda'$, we have

$$\frac{d}{dk} H(\lambda, k, t) = \frac{d}{dk} \int_0^\lambda h(\tilde{\lambda}, k, t) d\tilde{\lambda} = \int_0^\lambda \frac{d}{dk} h(\tilde{\lambda}, k, t) d\tilde{\lambda} < 0$$

since $\frac{d}{dk} h(\tilde{\lambda}, k, t) < 0$ for $\tilde{\lambda} < \lambda'$,

For $\lambda > \lambda'$,we have

$$\begin{aligned} \frac{d}{dk} H(\lambda, k, t) &= \frac{d}{dk} \int_0^\lambda h(\tilde{\lambda}, k, t) d\tilde{\lambda} = \frac{d}{dk} \int_0^\infty h(\tilde{\lambda}, k, t) d\tilde{\lambda} - \frac{d}{dk} \int_\lambda^\infty h(\tilde{\lambda}, k, t) d\tilde{\lambda} \\ &= 0 - \int_\lambda^\infty \frac{d}{dk} h(\tilde{\lambda}, k, t) d\tilde{\lambda} < 0 \end{aligned}$$

since $\frac{d}{dk} h(\tilde{\lambda}, k, t) < 0$ for $\lambda > \lambda'$. In the last line we used the fact that $\int_0^\infty h(\tilde{\lambda}, k, t) d\lambda = 1$ implies $\frac{d}{dk} \int_0^\infty h(\tilde{\lambda}, k, t) d\lambda = 0$.

Therefore, $\frac{d}{dk} H(\lambda, k, t) < 0$ for all (λ, k, t).

Proof of Theorem 7:

Part 1. We start with the Bellman equations for the discrete time approximation to the social planner's problem which are equations (90) and (91). From the structure of the optimization problem, the highest possible value of social welfare is $\frac{v}{r} - \underline{c}$ and the lowest possible value is 0. Let S be the space of functions defined on $(k, t, c) \in \Omega = \{Z^+ \times R^+ \times [\underline{c}, \overline{c}]\} \cup \{Z^+ \times R^+ \times \emptyset\}$ that are bounded above by $\frac{v}{r} - \underline{c}$, bounded below by 0 and continuous in c and t. Then S is a complete metric space under the sup norm.

We define the Bellman operator $T(\Delta t)$ on functions $W \in S$ as

$$\begin{aligned} T(\Delta t)W(k,t,c) &= \max\{\frac{v}{r} - c, T(\Delta t)W(k,t,\emptyset)\} \\ T(\Delta t)W(k,t,\emptyset) &= (1 - r\Delta t)\{m_1(k,t)\Delta t \int_{\underline{c}}^{\overline{c}} W(k+1, t+\Delta t, c) f(c) dc \\ &+ (1 - m_1(k,t)\Delta t) W(k, t+\Delta t, \emptyset)\} \end{aligned}$$

The fact that W takes values in $[0, \frac{v}{r} - \underline{c}]$ implies that $T(\Delta t)W$ does as well. That $T(\Delta t)W$ is a continuous function of t and c follows from standard theory given the structure of the choice set. Thus the Bellman operator maps S to S. It is straightforward to check that $T(\Delta t)$ is a contraction mapping and so has a unique fixed point $V^{\Delta t}$ in S. It also is straightforward to check that $T(\Delta t)$ maps functions that are increasing in k to functions that are increasing in k and maps functions that are decreasing in t and c to functions that are decreasing in t and c. This implies that $V^{\Delta t}$ is increasing in k and decreasing in t and c. As the unique fixed point of $T(\Delta t)$, $V^{\Delta t}$ is the unique solution to the Bellman equations (90) and (91).

We now derive the continuous time limit of Bellman equation (91). We use equation (90) to write the equation as

$$\begin{aligned} V(k,t,\emptyset) = (1 - r\Delta t)\{m_1(k,t)\Delta t \int_{\underline{c}}^{p(k+1,t+\Delta t)} (\frac{v}{r} - c) f(c) dc \\ + m_1(k,t)\Delta t(1 - F(p(k+1, t+\Delta t)))V(k+1, t+\Delta t, \emptyset) \\ + (1 - m_1(k,t)\Delta t)V(k, t+\Delta t, \emptyset)\} \end{aligned} \tag{102}$$

Subtracting $V(k,t,\emptyset)$ from both sides of equation (102) and dividing by Δt, we obtain

$$\begin{aligned} 0 = (1 - r\Delta t)m_1(k,t) \int_{\underline{c}}^{p(k+1,t+\Delta t)} (\frac{v}{r} - c) f(c) dc \\ + (1 - r\Delta t)m_1(k,t)(1 - F(p(k+1, t+\Delta t))V(k+1, t+\Delta t, \emptyset) \\ + \frac{V(k, t+\Delta t, \emptyset) - V(k,t,\emptyset)}{\Delta t} - V(k, t+\Delta t, \emptyset)(m_1(k,t) + r) \end{aligned} \tag{103}$$

Taking the limit as $\Delta t \to 0$, we obtain equation (95) which is

$$\begin{aligned} 0 = m_1(k,t)\Big\{ \int_{\underline{c}}^{p(k+1,t)} (\frac{v}{r} - c) f(c) dc + (1 - F(p(k+1,t))) V(k+1,t,\emptyset) \Big\} \\ + \frac{\partial V(k,t,\emptyset)}{\partial t} - V(k,t,\emptyset)(m_1(k,t) + r) \end{aligned} \tag{104}$$

Part 2. In this part, we analyze the optimal policy. Some properties of the optimal policy - that it is unique, decreasing in k and increasing in t - follow directly from equation (93) and the analogous properties of the value function. To derive an equation of motion for the optimal policy, we use equation (93) to obtain

$$\begin{aligned} \frac{\partial V(k,t,\emptyset)}{\partial t} &= -\frac{\partial p(k,t)}{\partial t} \\ V(k+1,t,\emptyset) &= \frac{v}{r} - p(k+1,t) \\ V(k,t,\emptyset) &= \frac{v}{r} - p(k,t) \end{aligned} \tag{105}$$

Substituting these expressions into equation (104) we express the Hamilton-Jacobi-Bellman equation in terms of the policy instead of the value function

$$\begin{aligned} \frac{\partial p(k,t)}{\partial t} = m_1(k,t)\Big\{ \int_{\underline{c}}^{p(k+1,t)} (\frac{v}{r} - c) f(c) dc + (1 - F(p(k+1,t)))(\frac{v}{r} - p(k+1,t)) \Big\} \\ - (\frac{v}{r} - p(k,t))(m_1(k,t) + r) \end{aligned} \tag{106}$$

Using integration by parts, we have

$$\int_{\underline{c}}^{p(k+1,t)} (\frac{v}{r} - c) f(c) dc = \frac{v}{r} F(p(k+1,t)) - p(k+1,t) F(p(k+1,t)) + \int_{\underline{c}}^{p(k+1,t)} F(c) dc \tag{107}$$

Substituting (107) into equation (106), we derive the equation of motion (96) which is

$$\frac{\partial p(k,t)}{\partial t} = m_1(k,t) \int_{\underline{c}}^{p(k+1,t)} F(c) dc + m_1(k,t)(p(k,t) - p(k+1,t)) - r(\frac{v}{r} - p(k,t)) \tag{108}$$

■

Proof of Theorem 8. The optimal policy $p(k,t)$ is the fixed point of a Bellman operator through its relationship to the value function given in equation (93). Below, we derive this Bellman operator. Let S be the space of functions defined on $(k,t) \in \Omega = \{Z^+ \times R^+\}$ that are bounded above by $\frac{v}{r} - \underline{c}$, bounded below by 0, and continuous in t. Then S is a complete metric space under the sup norm. The optimal policy $p(k,t)$ is the unique fixed point of the Bellman operator in S.

Define a subset S' of S consisting of functions g such that $g(k,t) \geq p^*(m_1(k,t))$ for all (k,t) where $p^*(\lambda)$ is the optimal policy when λ is known. To prove the theorem, we will show that the Bellman operator maps S' to itself, so that the fixed point $p(k,t)$ lies in S'.

First, we derive the Bellman operator for $p(k,t)$. We start with the Bellman equation (91). Instead of using the approximation $(1 - r\Delta t)$, we now use the exact discount rate in discrete time, which is $\frac{1}{1+r\Delta t}$. We write equation (91) as

$$V(k,t,\emptyset) = \frac{1}{1+r\Delta t}\{m_1(k,t)\Delta t \int_{\underline{c}}^{\overline{c}} V(k+1, t+\Delta t, c) f(c) dc + (1 - m_1(k,t)\Delta t) V(k, t+\Delta t, \emptyset)\} \tag{109}$$

Using equation (93), we write equation (109) in terms of the policy function as

$$\begin{aligned} \frac{v}{r} - p(k,t) = \frac{1}{1+r\Delta t}\{&m_1(k,t)\Delta t \int_{\underline{c}}^{p(k+1,t+\Delta t)} (\frac{v}{r} - c) f(c) dc \\ &+ m_1(k,t)\Delta t \int_{p(k+1,t+\Delta t)}^{\overline{c}} (\frac{v}{r} - p(k+1, t+\Delta t)) f(c) dc \\ &+ (1 - m_1(k,t)\Delta t) V(k, t+\Delta t, \emptyset)\} \end{aligned} \tag{110}$$

Using integration by parts, we obtain

$$\begin{aligned} \int_{\underline{c}}^{p(k+1,t+\Delta t)} (\frac{v}{r} - c) f(c) dc = &\frac{v}{r} F(p(k+1, t+\Delta t)) \\ &- p(k+1, t+\Delta t) F(p(k+1, t+\Delta t)) + \int_{\underline{c}}^{p(k+1,t+\Delta t)} F(c) dc \end{aligned} \tag{111}$$

Using (111), we have

$$\begin{aligned} &\int_{\underline{c}}^{p(k+1,t+\Delta t)} (\frac{v}{r} - c) f(c) dc + \int_{p(k+1,t+\Delta t)}^{\overline{c}} (\frac{v}{r} - p(k+1, t+\Delta t)) f(c) dc \\ &= \frac{v}{r} - p(k+1, t+\Delta t) + \int_{\underline{c}}^{p(k+1,t+\Delta t)} F(c) dc \end{aligned} \tag{112}$$

Substituting this into equation (110), we obtain

$$\begin{aligned} \frac{v}{r} - p(k,t) = &\frac{1}{1+r\Delta t} m_1(k,t)\Delta t \{(\frac{v}{r} - p(k+1, t+\Delta t)) + \int_{\underline{c}}^{p(k+1,t+\Delta t)} F(c) dc\} \\ &+ \frac{1}{1+r\Delta t}(1 - m_1(k,t)\Delta t)(\frac{v}{r} - p(k, t+\Delta t)) \end{aligned} \tag{113}$$

Rearranging, we obtain

$$\begin{aligned} p(k,t) = &\frac{v}{r}(1 - \frac{1}{1+r\Delta t}) + \frac{m_1(k,t)\Delta t}{1+r\Delta t} p(k+1, t+\Delta t) \\ &+ \frac{1}{1+r\Delta t}(1 - m_1(k,t)\Delta t) p(k, t+\Delta t) - \frac{m_1(k,t)\Delta t}{1+r\Delta t} \int_{\underline{c}}^{p(k+1,t+\Delta t)} F(c) dc \end{aligned} \tag{114}$$

We define the Bellman operator $T(\Delta t)$ on functions $g \in S$ as

$$T(\Delta t)g(k,t) = \frac{v}{r}(1 - \frac{1}{1+r\Delta t}) + \frac{m_1(k,t)\Delta t}{1+r\Delta t} g(k+1, t+\Delta t) \\ + \frac{1}{1+r\Delta t}(1 - m_1(k,t)\Delta t) g(k, t+\Delta t)) - \frac{m_1(k,t)\Delta t}{1+r\Delta t} \int_{\underline{c}}^{g(k+1,t+\Delta t)} F(c)dc \tag{115}$$

It is straightforward to show that $T(\Delta t)$ is a contraction mapping from S to S. By construction, the fixed point $p^{\Delta t}(k,t)$ is the optimal policy associated with the value function that solves the Bellman equations (90) and (109). In the limit as $\Delta t \to 0$, this policy must converge to the optimal policy $p(k,t)$ that satisfies the Hamilton-Jacobi-Bellman equations (94) and (95) and the equation of motion (96).

To finish the proof, we will show that $T(\Delta t)$ maps functions in S' to S'. This implies that $p(k,t) \in S'$ which is what we want to show. Let $g(k,t)$ be a function in S'. Then $g(k,t) \geq p^*(m_1(k,t))$ for all $(k,t) \in Z^+ \times R^+$.

Using equation (115) and the fact that $g(k, t+\Delta t) \geq p^*(m_1(k, t+\Delta t))$ we have

$$T(\Delta t)g(k,t) \geq \frac{v}{r}(1 - \frac{1}{1+r\Delta t}) + \frac{m_1(k,t)\Delta t}{1+r\Delta t} \big[g(k+1, t+\Delta t) - \int_{\underline{c}}^{g(k+1,t+\Delta t)} F(c)dc\big] \\ + \frac{1}{1+r\Delta t}(1 - m_1(k,t)\Delta t) p^*(k, t+\Delta t)) \tag{116}$$

The function

$$\Psi(x) = x - \int_{\underline{c}}^{x} F(c)dc \tag{117}$$

is increasing in c, which together with the fact that $g(k+1, t+\Delta t) \geq p^*(m_1(k+1, t+\Delta t))$, implies

$$T(\Delta t)g(k,t) \geq \frac{v}{r}(1 - \frac{1}{1+r\Delta t}) + \frac{m_1(k,t)\Delta t}{1+r\Delta t} \big[p^*(m_1(k+1, t+\Delta t)) - \int_{\underline{c}}^{p^*(m_1(k+1,t+\Delta t))} F(c)dc\big] \\ + \frac{1}{1+r\Delta t}(1 - m_1(k,t)\Delta t) p^*(k, t+\Delta t)) \tag{118}$$

The social planner's beliefs satisfy the martingale property

$$m_1(k,t) = \alpha m_1(k+1, t+\Delta t) + (1-\alpha) m_1(k, t+\Delta t) \tag{119}$$

where $\alpha = m_1(k,t)\Delta t$. By assumption $p^*(\lambda)$ is convex, so this implies

$$p^*(m_1(k,t)) \leq m_1(k,t)\Delta t p^*(m_1(k+1, t+\Delta t)) + (1 - m_1(k,t)\Delta t) p^*(m_1(k, t+\Delta t)) \tag{120}$$

Substituting this into equation (118), we have

$$T(\Delta t)g(k,t) \geq \frac{v}{r}(1-\frac{1}{1+r\Delta t}) + \frac{1}{1+r\Delta t}p^*(m_1(k,t)) - \frac{m_1(k,t)\Delta t}{1+r\Delta t}\int_{\underline{c}}^{p^*(m_1(k+1,t+\Delta t))} F(c)dc \tag{121}$$

From Lemma 6, we have $m_1(k,t) \geq m_1(k,t+\Delta t)$. From the martingale property (119), we then have $m_1(k,t) \leq m_1(k+1,t+\Delta t)$. Since $p^*(\lambda)$ is decreasing in λ, this implies that $p^*(m_1(k,t)) \geq p^*(m_1(k+1,t+\Delta t))$. Replacing $p^*(m_1(k+1,t+\Delta t))$ with $p^*(m_1(k,t))$ as the upper limit of the integral in equation (121), we have

$$T(\Delta t)g(k,t) \geq \frac{v}{r}(1-\frac{1}{1+r\Delta t}) + \frac{1}{1+r\Delta t}p^*(m_1(k,t)) - \frac{m_1(k,t)\Delta t}{1+r\Delta t}\int_{\underline{c}}^{p^*(m_1(k,t))} F(c)dc \tag{122}$$

We can write this as

$$\begin{aligned} T(\Delta t)g(k,t) &\geq \frac{r\Delta t}{1+r\Delta t}\Big\{\frac{v}{r} + \frac{p^*(m_1(k,t))}{r\Delta t} - \frac{m_1(k,t)\Delta t}{r\Delta t}\int_{\underline{c}}^{p^*(m_1(k,t))} F(c)dc\Big\} \\ &= \frac{r\Delta t}{1+r\Delta t}\Big\{\frac{1+r\Delta t}{r\Delta t}p^*(m_1(k,t)) + \frac{v}{r} - p^*(m_1(k,t)) - \frac{m_1(k,t)\Delta t}{r\Delta t}\int_{\underline{c}}^{p^*(m_1(k,t))} F(c)dc\Big\} \end{aligned} \tag{123}$$

Recall the result in equation (8) of the paper that $p^*(\lambda)$ satisfies

$$\frac{v}{r} - p^* = \frac{\lambda}{r}\int_{\underline{c}}^{p^*} F(c)dc \tag{124}$$

From this, we see that the last three terms in equation (125) cancel so that we have

$$T(\Delta t)g(k,t) \geq p^*(m_1(k,t)) \tag{125}$$

which is what we wanted to show. ■