Knowledge Creation and Diffusion with Limited Appropriation

Hugo Hopenhayn* UCLA

Liyan Shi† EIEF

February 2020

Abstract

Innovation is central to economic growth, but so is the diffusion of new knowledge. Such is the finding of recent macro papers that model the interaction between these two forces. Absent in this literature are three key elements that are the focus of this paper. First, we consider the role of frictions in matching innovators and imitators mediating the process of knowledge transmission. Second, we introduce the possibility of creative destruction, upon which event the innovator is replaced by the imitator. Third, while most of the recent literature has focused on the case where all surplus from knowledge transmission is captured by the imitators, we consider all ranges of possible shares that the innovators and imitators can appropriate and their impact on growth. In a simple one period model, we derive a modified Hosios condition for the optimal share when firms are ex-ante homogeneous. But we also find that as the degree of heterogeneity increases, the share of innovators must decrease to maximize growth. Our calibrated dynamic model suggests that the optimal share of surplus innovators appropriate should be in the medium range.

Keywords: Innovation, knowledge diffusion, endogenous growth, appropriation

*University of California, Los Angeles, email: `hopen@econ.ucla.edu`

†Einaudi Institute for Economics and Finance, email: `liyan.shi@eief.it`

1 Introduction

Knowledge creation and diffusion are the main forces behind economic growth. Starting with Lucas (2009), a series of recent papers in the macro literature have emphasized the role of diffusion as a contagion process where lagging firms learn the knowledge of more advanced firms in random matching. Moll and Lucas (2014) and Perla and Tonetti (2014) are examples. It is so assumed that the surplus generated is appropriated completely by the firm on the receiving end. In contrast, many forms of knowledge transmission involve some degree of appropriation by the transferring firms, such as technology licensing deals. Our paper captures this feature by modeling knowledge transfer as a bargaining problem between these two parties. More explicitly, we consider the role of the Nash bargaining weights in this problem and their impact on knowledge creation and transmission.

Intellectual property rights are a critical part of innovation policy. There is a long literature arguing in favor or against strong patent rights, trading off incentives for innovation and its costs. In our setting, stronger patent rights are associated with a higher bargaining weight for innovators. Our paper considers the impact this has on both, incentives for innovation and learning and its overall effect on economic growth. We explore this theoretically and quantitatively calibrating a model of innovation and knowledge diffusion.

An increase in the bargaining weight of the innovator, and thus its share of the surplus, has two opposing effects: 1) it directly encourages innovation; and 2) it discourages learning. In the presence of congestion a la Mortensen and Pissarides (1994), this in turn reduces the contact rate for innovators (increases for imitators), having a negative effect on innovation. At one extreme, when all surplus is appropriated by the firm transferring knowledge, thus holding up the learning firm, there are no incentives for learning and knowledge transfer disappears. At the other extreme, when the learning firm appropriates all surplus, innovation occurs only for its direct productive benefit to the innovator who disregards the value created by knowledge transfer.

Intellectual property rights play a dual role. On the one hand, they allow innovators to appropriate some of the rents from follow up firms that use or build up on these innovations. On the other hand, they can discourage competing innovations that follow up and destroy some of the rents obtained by the original innovator, i.e. creative destruction. The relative importance of each of these channels depends on the share of creative destruction as a fraction of total innovation. As we find, this plays an important role when considering the impact of intellectual property rights on economic growth.

In a simple one period model when all firms are ex-ante identical we show that the maximum level of growth is achieved at an intermediate bargaining weight for innovators that is the one suggested by the well-known Hosios condition (Hosios (1990)) in models of random matching.[1]When firms differ in their initial level of productivity, those above a certain threshold choose to innovate while those below choose to learn from the innovators. We find that as ex-ante heterogeneity increases, the optimal bargaining weight for innovators decreases and becomes zero when heterogeneity is sufficiently high. We also find that the optimal bargaining weight increases with the share of creative destruction, measured by the extent to which adoption by a learning firm decreases the original innovator's value.

Heterogeneity matters for two reasons. Firstly, it corresponds to a component of knowledge that is exogenous to innovation efforts of a firm. When the importance of this component overwhelms the one resulting from innovation effort, stronger IP modeled here as higher innovator share of surplus gives rents to innovators without affecting much their innovation or participation decisions, while discouraging imitation and learning at the same time. Secondly, with ex-ante heterogeneity the marginal type (the one at the threshold) represents a less valuable match for imitators to learn from. Thus, the positive external effect on imitators is smaller, while the negative crowding out effect on innovators is the same.

We explore quantitatively the question in a dynamic general equilibrium scenario that is a small variation of the basic setup in Benhabib, Perla and Tonetti (2017) with the addition of matching congestion. There is a fixed set of agents that can either operate as productive firms or learn from others. Firm's productivity follows a Brownian motion where drift is a function of costly innovation effort and volatility is given. Learning agents contact a randomly selected firm at a rate that is determined via a matching function. Upon meeting, the two agents split the value of the learning firm according to Nash bargaining. With some probability, the original innovator loses the value of its innovation and goes back to join the pool of learning agents. This parameter measures the relative share of creative destruction.

The model is calibrated to match a series of aggregate moments.[2] Given parameter values, we calculate the share of surplus that maximizes growth. As in our simple model, we find that as volatility becomes large and heterogeneity increases, the optimal innovator share of surplus decreases. In our preliminary calibration, where volatility is chosen to match moments of the size distribution of firms, the optimal share is in the medium range. Given

[1]The intuition for this result is as follows. Holding fixed the innovation decision μ, the Hosios condition guarantees that the equilibrium delivers the optimal fraction of innovators in the population, i.e. the one that maximizes total surplus. What is more surprising, is that total surplus when also taking into account the equilibrium choice of μ is also maximized at this point. The intuition behind this result is that in equilibrium innovator profits from innovation are proportional to total surplus, when surplus is maximized so are the total (direct and indirect) incentives for innovation.

[2]These moments include: aggregate growth rate, interest rate, size distribution of firms, volatility of firm size, share of creative destruction (as given by Garcia-Macia, Hsieh and Klenow (2019)) and the ratio of public to private returns to innovation (as given by Bloom, Schankerman and Van Reenen (2013)).

these parameter values, total growth falls with innovator share mostly because of the severe drop in knowledge transmission, as the fraction of learning firms decreases dramatically with the bargaining weight of innovators. The optimal fraction of innovators is also sensitive to parameters in the matching function. As we vary the elasticity of the matching function from zero to one, the optimal fraction of innovators also goes from zero to one. This shows the importance of matching frictions in the process of knowledge transfer when considering the optimal assignment of intellectual property rights.

1.1 Related Literature

Our paper builds on several strands of literature, including the papers on knowledge diffusion, patent policy and innovation, and matching frictions, as well as recent papers attempting to measure returns to innovation and the extent of creative destruction.

Kortum (1997) considers a setting where firms sample from a fixed distribution of ideas, recognizing that a key to sustained exponential growth is that the stationary part of this distribution has a Pareto upper tail. In Luttmer (2007), this distribution is endogenously generated in the steady state, as entrants learn from the distribution of incumbent's productivity. A similar mechanism is developed in Lucas (2009) where all firms exogenously learn from others, where the initial distribution of knowledge has a Pareto tail.[3] Moll and Lucas (2014) and Perla and Tonetti (2014) endogenize the allocation of resources devoted to learning. In these models, resources can be allocated either to production or learning.

Our model builds on König, Lorenz and Zilibotti (2016) and Benhabib et al. (2017), where firms make an optimal choice between innovation and imitation. In a recent addition to their paper (contemporaneous to ours), Benhabib et al. (2017) consider briefly the role of bargaining. Our analysis differs in several respects: First, we focus on the question of intellectual property protection and derive explicitly the bargaining weights from the strength of patent enforcement. Second , we focus on the relationship between patent enforcement and growth, and how the growth maximizing policy is affected by several characteristics of the environment, including innovator heterogeneity, matching frictions, and creative destruction. The latter two characteristics are new features we introduce and absent in their model.

We borrow from recent papers that quantify the technology and product market spillover effects of innovation. Mapping our model to their framework, we use Bloom et al. (2013)'s calculation of the social versus private returns to innovation. Similarly, we use the importance of creative destruction measured by Garcia-Macia et al. (2019).

[3] A class of hybrid models where firm productivity can evolve from the sampling of an exogenous distribution (as in Kortum (1997)) and the distribution of other firms' knowledge (as in Luttmer (2007) and Lucas (2009)) can be found in Alvarez, Buera and Lucas (2008) and Buera and Oberfield (Forthcoming).

Our insights of the optimal bargaining weights relate to the literature on matching frictions in random search. Using an alternative protocol as opposed to Nash bargaining, our expression for the optimal weights resembles Hosios (1990)'s condition, by adjusting for the creative destruction events. In the presence of heterogeneity, however, this condition fails as pointed out by Shimer and Smith (2001).

2 One-Period Model

This section provides a simple one-period model to illustrate how the bargaining weight trades off incentives for innovation versus imitation, considering the possibilities of replacement of innovators by imitators as well as congestion in knowledge diffusion. We first study the case in which firms are homogenous before moving on to heterogeneous firms.

2.1 Homogenous Firms

The economy lasts for one period and consists of a unit mass of firms. We consider a pure endowment economy, where the firm's output is equal to its productivity. The firms produce identical goods. They are initially endowed with the same productivity $z = 1$, which can be improved through either innovation or imitation.

Timing of events. A two-stage game takes place. In the first stage, firms choose whether to innovate or wait to imitate. If they decide to innovate, they can improve their productivity to μz at a cost of $c(\mu) z$. The innovation cost function is convex and strictly increasing in the innovation intensity μ. The proportion of firms that invest in innovation is denoted by α, and the remaining $1 - \alpha$ proportion wait to imitate.

In the second stage, the innovators and the imitators search and match randomly in the market for knowledge transfer. The market is two-sided, with innovators and imitators searching on each side. The aggregate number of matches is given by a matching function $M(\alpha, 1 - \alpha)$, which is assumed to be homogenous of degree one, once differentiable, and increasing in both arguments. Upon being matched, the imitator can acquire the innovator's technology through a knowledge transfer by paying a licensing fee.[4] With probability $\theta \in [0, 1]$ the knowledge transfer results in a creative destruction of the innovator firm, in which event the innovator is replaced by the imitator. To determine the amount of licensing fee, they bargain over the imitator's surplus from the knowledge transfer. The bargaining weight

[4]We assume that the imitator fully learns the innovator's knowledge . One could generalize it to a setting of limited learning, as in the mergers and acquisition model by David (2017). The newly acquired knowledge depends on not only the innovator's level but also the imitator's own level.

for the transferring party is $\beta \in [0, 1]$. That is, the transferring firm appropriates β share of the surplus and the learning firm appropriates $1 - \beta$ share.

At the end of second stage, the firms produce. Firms that have acquired the new technology produce output μ and the rest of the firms produce zero output.

Matching and congestion. The matching function in the market for knowledge transfer introduces the possibility of congestion in learning. If there are more imitators searching for better technologies to copy, it becomes more difficult for them to find such opportunities. To illustrate why congestion matters, consider a special case in which the probability of meeting an innovator is constant, and hence independent of the composition of imitators and innovators, as in the knowledge diffusion models by Perla and Tonetti (2014) and Buera and Oberfield (Forthcoming). For example, if the matching function is $M(\alpha, 1 - \alpha) = \phi(1 - \alpha)$, there is no congestion on the imitator's side, while the innovator's side is fully congested. An infinitesimally small amount of innovators are needed to spread the knowledge to the whole economy.

Bargaining weights. The innovator's bargaining weight represents the enforcement probability under the intellectual property rights regime. This interpretation can be illustrated through a simple game, which goes as follows. Upon being matched, the innovator makes a take-it-or-leave-it offer to the imitator, asking for a licensing fee t for the technology transfer. If the imitator does not accept and uses the technology without a license, there is a probability β of being caught. Once caught, the imitator is then

not allowed to use the innovator's technology. Suppose the imitator's surplus from the technology transfer is S. The imitator has an incentive-compatibility constraint:

$$S - t \geq (1 - \beta) S,$$

which implies that the innovator would ask for $t = \beta S$ and the imitator would accept. Stronger intellectual property rights protection in favor of innovators leads to stronger bargaining power of innovators, which allows them to appropriate more rents from knowledge diffusion.[5]

[5]The bargaining protocol differ from the alternating offers in Nash bargaining. With Nash bargaining, the innovator and imitator would split the social surplus from the knowledge transfer. In contrast, in the enforcement threat game here, the firms split the imitator's surplus from knowledge transfer.

Competitive equilibrium. In the decentralized economy, in stage one, the innovators choose innovation intensity that maximizes their net value:

$$\max_{\mu} \quad \left[1 + (\beta - \theta) \frac{M(\alpha, 1-\alpha)}{\alpha}\right] \mu - c(\mu),$$

where the payoff of innovation comes in two parts: a direct payoff through increased output, μ, and an indirect payoff through knowledge transfer net of the losses from the creative destruction, $(\beta - \theta) \frac{M}{\alpha} \mu$. The optimality condition for the innovation intensity is such that the marginal private payoff equals marginal cost,

$$1 + (\beta - \theta) \frac{M(\alpha, 1-\alpha)}{\alpha} = c'(\mu). \tag{1}$$

As firms are ex-ante identical, we obtain an indifference condition where the values of innovators and non-innovators are equal. That is, the value appropriated by the imitators equals the combined payoff from innovation minus the cost of innovation:

$$(1 - \beta) \frac{M(\alpha, 1-\alpha)}{1-\alpha} \mu = \left[1 + (\beta - \theta) \frac{M(\alpha, 1-\alpha)}{\alpha}\right] \mu - c(\mu). \tag{2}$$

Optimal appropriation. The bargaining weight plays a key role in trading off the incentives for innovation and diffusion. First, as characterized in equation (1), the bargaining weight exerts a direct holdup effect on investment through the extent of knowledge appropriability, or the intensive margin of innovation. A higher bargaining weight of innovators encourages innovation and discourages learning. More drastically, at one extreme, when all surplus is appropriated by the firm transferring knowledge, thus holding up the learning firm, there are no incentives for learning and knowledge transfer disappears; at the other extreme, when the learning firm appropriates all surplus, innovation occurs only for its direct productive benefit to the innovator who disregards the value created by knowledge transfer. Second, in equation (2), the bargaining weight indirectly affects the equilibrium fraction of innovators and imitators and thus the matching rates of both groups, or the extensive margin of innovation and diffusion. A higher β increases the fraction of innovators and thus reduces the contact rate for innovators and increases that for imitators. The second effect is similar to the standard appropriability concern and congestion externality of random search in the labor market search literature.

The investment and matching aspects described above imply that the equilibrium allocation is suboptimal compared to the first-best. When considering the incentives for innovation alone, the innovator should appropriate the entire transfer surplus $\beta = 1$, as it aligns the innovator's marginal private value of innovation to the marginal social value. When considering the external effects on matching alone, the bargaining weight should be such that the external congestion effects are equalized. As is well known from Hosios (1990), this is achieved when innovating firms appropriate a share of the transfer surplus that is equal to the elasticity of the matching function with respect to the fraction of innovators. It follows that the assignment of bargaining weights cannot simultaneously align the private values of investment in innovation and search with their social values.

It is true that with a sufficiently rich set of instruments, such as direct subsidies to innovation and strength of patent protection (interpreted here as the bargaining weights), the first-best allocation can be supported. However, as discussed in the patent literature, direct subsidies might be difficult to implement when innovation is not directly observed. Following this literature, we ask in this paper what can be achieved when the planner has at its disposal only the assignment of bargaining weights. These weights β are chosen to maximize net output:

$$Y = \alpha\left[\mu - c\left(\mu\right)\right] + M\left(1-\theta\right)\mu, \tag{3}$$

subject to innovation and imitation decisions, μ and α, satisfying firms' investment optimality condition (1) and the indifference condition (2).

Proposition 1. *With ex-ante homogenous firms, the optimal bargaining weight is,*

$$\beta^{*} = \left(1-\theta\right)M_1\frac{\alpha}{M} + \theta. \tag{4}$$

The optimal bargaining weight in equation (4) resembles Hosios (1990)'s condition, by adjusting for the creative destruction probability. The intuition for this result is as follows. Holding fixed the innovation decision μ, the condition guarantees that the equilibrium delivers the optimal fraction of innovators in the population, i.e. the one that maximizes the amount of knowledge transfer $M\left(\alpha, 1-\alpha\right)$. More surprisingly, when also taking into account the equilibrium choice of μ, the total surplus is also maximized at this point. This is because in equilibrium the innovator's profit from innovation is proportional to total surplus, when surplus is maximized so are the total (direct and indirect) incentives for innovation.

Introducing matching congestion in knowledge diffusion matters and is crucial for these results. Again it is instructive to consider an economy with no congestion on the imitator's side where the chance of finding an innovator is independent of the ratio of innovators to imitators, so the total amount of match is $M\left(\alpha, 1-\alpha\right) = \phi\left(1-\alpha\right)$.[6] According to equation (4), the optimal bargaining weight β^{*} would be approaching θ, just compensating

[6] This special case corresponds to the knowledge diffusion process in Perla and Tonetti (2014), Perla, Tonetti and Waugh (2015), and Benhabib et al. (2017).

the innovators for the losses from creative destruction. The intuition is obvious: innovation is costly, yet only an infinitesimally small fraction of innovators are needed to spread the knowledge to the whole economy.

2.2 Heterogenous Firms

To account for the existence of substantial knowledge difference in the economy, we extend the one-period model to include ex-ante heterogeneity in knowledge. This will also help us to interpret our subsequent results in the dynamic model, where the heterogeneous level of firm productivity are endogenous outcome of innovation and diffusion. We assume that the unit mass of firms have initial productivity z, following a distribution with cumulative density function $F(z)$.

In equilibrium, there is selection of more productive firms into innovation activities and less productive ones into imitation. The productivity of the marginal firm who is indifferent between innovating or copying is denoted by productivity z^*. The fraction of innovators is $\alpha = 1 - F(z^*)$. The marginal firm is characterized by the indifference condition:

$$(1-\beta)\frac{M(\alpha, 1-\alpha)}{1-\alpha}\int_{z^*} z dF(z)\,\mu = \left\{\left[1 + (\beta-\theta)\frac{M(\alpha, 1-\alpha)}{\alpha}\right]\mu - c(\mu)\right\} z^*. \tag{5}$$

The net output is:

$$Y = \int_{z^*} z dF(z)\left[\mu - c(\mu)\right] + M(1-\theta)\frac{1}{\alpha}\int_{z^*} z dF(z)\,\mu.$$

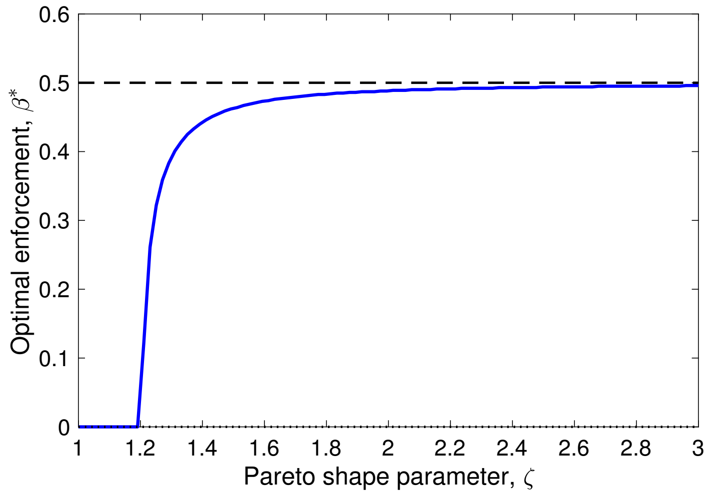
Note that the social surplus from knowledge diffusion per match depends not only on the extent of investment μ but also the average innovator $\frac{1}{\alpha}\int_{z^*} z dF(z)$.

Proposition 2. *With ex-ante heterogenous firms, the optimal bargaining weight is lower than with homogenous firms,*

$$\beta^* < (1-\theta) M_1 \frac{\alpha}{M} + \theta.$$

Heterogeneity contributes to a lower optimal bargaining weight because there is more social value from knowledge diffusion to be gathered. In addition to facilitating the diffusion of new knowledge created, a lower bargaining weight also helps to capture the gains from bridging the existing knowledge gaps, by having more imitators searching for better technologies. It can be identified to the component of knowledge diffusion exogenous to the innovation efforts of a firm. When the importance of this component overwhelms the one resulting from innovation effort, stronger intellectual property rights modeled here as higher

Figure 1: Optimal bargaining weight with heterogeneous firms



Notes: The matching function is $M(\alpha, 1-\alpha) = \phi\alpha^{\omega}(1-\alpha)^{\omega}$, with $\phi = 0.1$ and $\omega = 0.5$. The innovation cost function is $c(\mu) = \frac{\gamma}{\varepsilon+1}\mu^{\varepsilon+1}$, with $\gamma = 1$ and $\varepsilon = 1$. The creative destruction probability is $\theta = 2/3$.

β gives rents to innovators without affecting much their innovation or participation decisions, while discouraging imitation and learning at the same time.

To be precise, with ex-ante heterogeneity, the marginal firm (the one at the threshold) represents a less valuable match for imitators to learn from than the average innovating firms. Thus, the positive external effect on imitators is smaller, while the negative crowding out effect on innovators is the same. The Hosios condition for homogeneous firms balances out these two effects to get the optimal fraction of innovators. Since positive externalities are smaller with heterogeneity, less innovators will be optimal and this is induced in equilibrium by giving innovators a smaller bargaining weight.

Firms in the tail. To see exactly how heterogeneity changes the optimal bargaining weight, we construct one simple numerical example. This example also relates to the results in our dynamic model. The ex-ante knowledge level follows a Pareto distribution,

$$F(z) = 1 - \left(\frac{\underline{z}}{z}\right)^{\zeta}, \ \forall z \in [\underline{z}, \infty) \ \text{where } \zeta > 1.$$

The shape parameter ζ captures the extent of heterogeneity present in the economy. A higher ζ implies a thinner Pareto tail in productivity distribution and less heterogeneity. As ζ approaches infinity, heterogeneity would vanish and the economy reverts back to the ex-ante homogeneous case in the previous section.

The matching function has a constant elasticity, $\omega = M_1 \frac{\alpha}{M} = 0.5$. Figure 1 shows the level of optimal bargaining power that solves the planner's problem as we vary the extent of heterogeneity. Overall, more heterogeneity leads to lower optimal bargaining power. When there is sufficient heterogeneity, in this example ζ below somewhere close to 1, the optimal bargaining weight is zero. If the gain from diffusion of existing knowledge is large enough, as illustrated in the numerical example, the optimal bargaining weight of innovators is pushed all the way to zero. When heterogeneity vanishes, i.e. $\zeta \to \infty$, the optimal bargaining weight converges to the one implied by condition (4).

3 Dynamic Model

We now build a dynamic model which features the same ingredients of knowledge creation and diffusion as the one-period model. This dynamic model differs from the static one in the sense that the extent of knowledge heterogeneity in the economy is an endogenous outcome of innovation and diffusion. It also allows us to carry out quantitative assessment of the optimal level of appropriation.

3.1 Environment

[there is some redundancy in the discussion]

Time is infinite and continuous. The economy consists of a continuum measure-one of firms. The firms produce identical goods and are characterized by their level of productivity Z. We consider a pure endowment economy, where firm output is equal to its productivity.[7] The log productivity is denoted by $z \equiv \log(Z)$. At time t, the distribution of firms follows a cumulative density function $F(z,t)$. There is a representative household with preferences for consumption

$$\int_0^\infty e^{-\rho t} \log(C(t))\, dt.$$

The implied interest rate in this economy is $r(t) = \rho + C'(t)/C(t)$.

Firm productivity follows the stochastic process:

$$dz = i\,(\mu dt + \sigma(\mu)\, dB) + (1-i)\,(\tilde{z} - z)\, dJ, \tag{6}$$

where $i = 1$ indicates the decision to innovate and $i = 0$ the decision to imitate. An $\alpha(t)$ fraction of firms innovate. The contact rate for an innovator to meet an imitator is $q(\alpha(t)) \equiv M(1, (1-\alpha(t))/\alpha(t))$. The contact rate for imitators is $p(\alpha(t)) \equiv M(\alpha(t)/(1-\alpha(t)), 1)$. The innovating firm incurs a flow cost of innovation, $c(\mu)\exp(z)$, proportional to its productivity. The imitating firm has a Poisson rate of meeting firms with some superior technology

[7] If we were to introduce inputs other than technology in production, the results would not change.

$\tilde{z}$ and acquiring that technology, in which case its productivity jumps upwards, i.e. $dJ = 1$. With probability θ, a knowledge transfer results in a creative destruction event, upon which the innovator is replaced by the imitator.

Firm's endogenous innovation decision affects not only the expected level but also potentially the variance of productivity improvement. We capture the effect on the variance with a flexible specification of the Brownian motion, which has a standard deviation $\sigma(\mu)$ as a function of innovation intensity μ. While we maintain the general notation in the model, we study two particular innovation processes. The first innovation process is a *fixed-volatility* process, represented by a standard Brownian motion with a standard deviation σ. This model of innovation is commonly used in the literature, for example, in the continuous time model of Luttmer (2007) or, equivalently, in the discrete time model of **?**. We introduce an alternative *scalable-volatility* innovation process, represented by a Brownian motion with standard deviation $\sigma\sqrt{\mu}$. To micro-found this process, we consider firm innovation as experimentation through a set of risky projects. Let μ denote the number of projects carried out at a given point of time. Each project generates some incremental productivity improvement with outcome drawn independently from a normal distribution $N(1, \sigma^2)$. Together the μ projects generate productivity improvement drawn from a normal distribution $N(\mu, \sigma^2\mu)$.

3.2 Firm's Problem

Consider the problem of a firm with productivity z at time t. The firm decides whether to innovate or imitate, weighting the values it can obtain by innovating, $V^i(z,t)$, and by copying, $V^c(t)$:[8]

$$\max\left\{V^i(z,t), V^c(t)\right\}.$$

Since more productive firms has an advantage in and benefit more from transferring knowledge, there exists a threshold $\underline{z}(t)$ such that the firms above the threshold choose to innovate and those below the threshold imitate. The innovation threshold $\underline{z}(t)$ satisfies the following value matching and smooth pasting conditions:

$$V^c(t) = V^i(\underline{z}(t), t), \tag{7}$$

$$0 = V^i_z(\underline{z}(t), t). \tag{8}$$

[8] By assuming that the imitators are idle in production, their value function doesn't depend on their knowledge level. Conversely, if we were to assume that imitators also produce, their value function will depend on their knowledge level; the marginal innovator is different from the average innovator. We discuss later that, quantitatively, this distinction is important.

The value matching condition is the indifference condition such that the firms at the innovation threshold obtain the same value from either innovating or copying. The smooth pasting condition is necessary here since firms at the innovation threshold are moving backward relative to the threshold over time [better argument?]. The measure of innovators satisfies

$$\alpha(t) = 1 - F(\underline{z}(t), t). \tag{9}$$

The innovating firm's value function, $V^i(z,t)$, satisfies the Hamilton–Jacobi–Bellman (HJB) equation:

$$\begin{aligned} r(t)V^i(z,t) = \max_{\mu} \Big\{ & \exp(z) + q(\alpha(t))(\beta - \theta)\big(V^i(z,t) - V^c(t)\big) \\ & -c(\mu)\exp(z) + \mu V^i_z(z,t) + \frac{1}{2}\sigma(\mu)^2 V^i_{zz}(z,t) + V^i_t(z,t) \Big\}. \end{aligned} \tag{10}$$

In equation (10), the flow payoff to the innovating firm on the right-hand side of the first line consists of two parts: a direct payoff from production profit $\exp(z)$ and an indirect payoff through knowledge transfer. The expected flow payoff from knowledge transfer is β fraction of the imitator's surplus from knowledge transfer, taking into the probability of meeting an imitator $q(\alpha(t))$, net of the change in its value if the innovator is replaced. The first three terms in the second line capture the innovation cost and the change in value due to productivity innovations. The innovation intensity decision rule $\mu(z,t)$ is such that the marginal cost of innovation is equal to the marginal benefit from productivity improvement:

$$c'(\mu(z,t))\exp(z) = V^i_z(z,t) + \sigma(\mu(z,t))\sigma'(\mu(z,t))V^i_{zz}(z,t). \tag{11}$$

The imitating firm's value function, $V^c(z,t)$, satisfies the HJB equation:

$$r(t)V^c(t) = p(\alpha(t))(1-\beta)\big(\mathbb{E}\left[V^i(\tilde{z},t)\,|\tilde{z} > z(t)\right] - V^c(t)\big) + V^c_t(t). \tag{12}$$

In equation (12), the flow payoff also consists of two parts: a direct payoff from profit $\exp(z)$ and an indirect payoff from retained surplus after paying licensing fees. The payoff from imitation is equal to the probability of meeting an innovator, $p(\alpha(t))$, times the $1-\beta$ fraction of the imitator's surplus upon a meeting.

3.3 Equilibrium Definition

The distribution of firm productivity evolves over time according to the following Kolmogorov Forward (KF) equation: $\forall z \geq \underline{z}(t)$,

$$f_t(z,t) = q(\alpha(t))(1-\theta) f(z,t) - \frac{\partial(\mu(z,t) f(z,t))}{\partial z} + \frac{1}{2}\frac{\partial^2 \left(\sigma(\mu(z,t))^2 f(z,t)\right)}{\partial z^2}. \quad (13)$$

This distribution is shaped by forces of innovation and diffusion. On the one hand, there is an inflow of firms from the left to the right of the innovation threshold, as they acquire superior technology and jump ahead in the knowledge scale. On the other hand, the innovating firms are advancing along the knowledge scale at the speed of their innovation intensity $\mu(z,t)$.

Definition 1 (Competitive Equilibrium)**.** A competitive equilibrium consists of value functions $\{V^i(z,t),\ V^c(t)\}$, innovation decision rule $\mu(z,t)$, innovation threshold $\underline{z}(t)$, fraction of innovators $\alpha(t)$, and productivity distribution $F(z,t)$, given the initial productivity distribution $F(z,0)$, such that equations (7) to (13) are satisfied.

Lemma 1 (Innovation Intensity)**.** *Firm innovation intensity $\mu(z,t)$ increases in the level of productivity z and converges to an upper bound*

$$\lim_{z\to\infty} \mu(z,t) = \bar{\mu}(t).$$

This lemma follows from the observation that the firm's problem is an optimal stopping problem. The further away the innovator is from the innovation threshold $\underline{z}(t)$, the longer it is expected to benefit from improved productivity before switching to the side of imitators. Hence, the innovation intensity $\mu(z,t)$ increases with the innovator's productivity z. At the extreme, for firms on the knowledge frontier, i.e., $z \to \infty$, the distance to the innovation threshold becomes irrelevant, and the optimal stopping point is no longer a concern. This feature is also reflected in the HJB equation (10) for the innovator's value function: the nonlinear part due to the outside option of the imitating firms, $V^c(t)$, becomes infinitely small in comparison to the direct payoff.

3.4 Balanced Growth Path

We focus on the balanced growth path (BGP) along which the economy grows at a constant rate g. The aggregate consumption $C'(t)/C(t) = g$. The value functions of innovators and imitators satisfy

$$V^i(z+gt,t) = \exp(gt)\, V^i(z) \tag{14}$$
$$V^c(z+gt,t) = \exp(gt)\, V^c(z), \tag{15}$$

and the productivity distribution can be transformed into a stationary one,

$$f(z+gt,t) = f(z). \tag{16}$$

The innovation threshold also grows at rate g, $\underline{z}(t) = \exp(gt)\,\underline{z}$. The innovation intensity converges to a constant upper bound, $\bar{\mu}(t) = \bar{\mu}$.

Assumption 1. *The preference and technology parameters satisfy*

$$\rho > -\sigma(\bar{\mu})\sqrt{2q(\alpha)} + (\beta-\theta)\,q(\alpha) + \frac{1}{2}\sigma(\bar{\mu})^2.$$

Assumption 1 states that agents discount future consumption at a sufficiently high rate. This assumption is merely to ensure that, along the BGP, firm value functions are bounded and hence well-defined. In the quantitative section, the assumption is guaranteed given that the contact rate $q(\alpha)$ and the standard deviation $\sigma(\bar{\mu})$ are small in magnitude.

Assumption 2. *The initial productivity distribution $F(z,0)$ has a bounded support.*

Under this assumption, the initial set of knowledge in the economy is bounded and limited. Therefore creation of new knowledge through innovation is essential to generate sustained economic growth. Otherwise, absent any innovative effort, the economy would eventually converge to the highest productivity level in the initial economy as a consequence of knowledge diffusion.

On a technical level, this assumption is a sufficient condition to ensure the existence of a unique BGP, as in Luttmer (2007). Otherwise there can potentially exist a continuum of equilibria associated with different growth rates and stationary distributions. For example, Benhabib et al. (2017) distinguishes between a unique growth path and hysteresis, depending on whether the initial productivity distribution has a bounded support. Here we consider an economy with a unique GBP.

Proposition 3 (BGP)**.** *Under Assumption 1 and 2, there exists a unique BGP.*

1. Along the BGP, all variables grow at rate

$$g = \bar{\mu} + \sigma(\bar{\mu})\sqrt{2q(\alpha)(1-\theta)}, \tag{17}$$

where the upper bound of innovation $\bar{\mu}$ is characterized by

$$c'(\bar{\mu}) = \frac{1 + \sigma(\bar{\mu})\,\sigma'(\bar{\mu})}{\rho + g - q(\alpha)(\beta - \theta) - \bar{\mu} - \frac{1}{2}\sigma(\bar{\mu})^2}\,(1 - c(\bar{\mu}))\,; \tag{18}$$

2. The endogenous stationary distribution, $\forall z > \underline{z}$, has a Pareto tail with the shape parameter

$$\zeta = \frac{\sqrt{2q(\alpha)(1-\theta)}}{\sigma(\bar{\mu})}. \tag{19}$$

Growth decomposition. The first part of Proposition 3 provides a formula for the growth rate. Equation (17) decomposes the growth rate into two sources: the speed of innovation by firms on the knowledge frontier, $\bar{\mu}$, and the extent of knowledge diffusion from innovating to imitating firms, $\sigma(\bar{\mu})\sqrt{2q(\alpha)(1-\theta)}$. Similar formulas are obtained in continuous-time models in previous paper by Luttmer (2007, 2012) and Benhabib et al. (2017). A new feature here is that each component in the formula is an endogenous outcome driven by the incentives for innovation and imitation, including the fraction of firms innovating α, correspondingly the matching rate $q(\alpha)$, and the innovation intensity $\bar{\mu}$. The relative strength of the bargaining power between innovators and imitators affects their incentives and therefore trades off the two sources of growth.

As the bargaining position of innovators strengthens, directly, the innovating firms appropriate more rents from knowledge diffusion, which induces a higher innovation effort. This is demonstrated by the optimality condition (18) that characterizes the upper bound of innovation $\bar{\mu}$. At the same time, an indirect force due to congestion offsets that innovation incentive. As a larger proportion of firms decide to innovate rather than imitate, the diffusion rate $q(\alpha)$ goes down. The indirect force could dominate the positive direct effect. At the extreme, when the innovators have all the bargaining power, i.e., $\beta = 1$, no firm would ever want to be on the imitation side, and knowledge diffusion completely disappears, $q(\alpha) = 0$. Combining the two effects, the indirect payoff to innovation from knowledge transfer $(\beta - \theta)\,q(\alpha)$ exhibits a hump shape in the innovator's bargaining weight β.

The overall effect of the bargaining weights on the growth rate is most likely non-monotonic. In the lower range of the innovator's bargaining weight, increased growth through innovation could offset decreased diffusion. In the upper range, the growth rate is clearly decreasing in the innovator's bargaining weight.

Productivity heterogeneity. The second part of Proposition 3 characterizes the detrended stationary productivity distribution. The distribution of innovators does not have an analytical solution, due to the varying level of innovation intensity depending on the firm's distance to the innovation threshold. However, on the knowledge frontier, as the innovation intensity converges to its constant upper bound, we can characterize the distribution in the right tail. The productivity of innovating firms follows an asymptotic Pareto distribution with a shape parameter in equation (19). The shape of the right tail is a useful measure of productivity heterogeneity. The randomness in innovation outcome and the diffusion rate together determine shape of productivity distribution in the right tail. A higher standard deviation $\sigma(\bar{\mu})$ leads to a fatter right tail, while a higher diffusion rate $q(\alpha)$ leads to a thiner right tail.

Innovation and diffusion are two opposing forces in shaping the extent of equilibrium heterogeneity. On the one hand, innovation contributes to firm heterogeneity, stretching out the productivity distribution. The dispersion is mainly due to the stochastic nature of innovation outcome, as captured by the standard deviation of the Brownian motion. It is particularly true in our scalable-volatility specification of the innovation process, $\sigma(\bar{\mu}) = \sigma\sqrt{\bar{\mu}}$, where a higher innovation intensity through more experimentation leads to more dispersed outcome. Knowledge diffusion, on the other hand, compresses the productivity distribution as the less productive firms catch up with the highly-productive firms.

Lemma 2 (Firm Growth)**.** *For the innovating firms, the expected growth rate and the variance of the growth rate converge to*

$$\lim_{z\to\infty} \mathbb{E}\left[z_{t+\tau} - z_t\right] = \bar{\mu}\tau$$

$$\lim_{z\to\infty} var\left(z_{t+\tau} - z_t - g\tau\right) = \sigma\left(\bar{\mu}\right)^2 \tau.$$

The intuition for this lemma is straightforward. The productivity of innovating firms follows a stochastic diffusion process. For the firms on the knowledge frontier, i.e., $z \to \infty$, the expected growth rate is equal to the upward drift, the volatility of the growth rate is equal to the standard deviation.

3.5 Social Value of Innovation

We derive the social value of innovation, taking into account the incremental value created through future knowledge transfers. Let $W^i(z,t)$ denote the social value associated with an innovator of productivity z at time t. On the balanced path, $W^i(z+gt,t) = \exp(gt)\, W^i(z)$. On the knowledge frontier, $\lim_{z\to\infty} W^i(z)/\exp(z) = W^i$.

Proposition 4 (Social Value)**.** *The marginal social value of innovation converges to*

$$\lim_{z\to\infty} W^i = \frac{1+\sigma(\bar{\mu})\,\sigma'(\bar{\mu})}{\rho+g-q(\alpha)(1-\theta)-\bar{\mu}-\frac{1}{2}\sigma(\bar{\mu})^2}\left(1-c(\bar{\mu})\right). \tag{20}$$

Table 1: Calibrated parameters

Parameter	Value		Moment
	Fixed Vol.	Scalable Vol.	
Discount rate ρ	0.02	0.02	interest rate & growth rate
Standard deviation σ	0.08	0.94	vol. of large firm growth
Creative destruction prob. θ	2/3	2/3	Garcia-Macia et al. (2019)
Bargaining weight β	0.3	0.3	R&D returns by Bloom et al. (2013)
Matching function efficiency λ	0.034	0.038	diffusion rate
Matching function curvature ω	0.5	0.5	preset
Innovation cost elasticity ε	1	1	preset
Innovation cost scale γ	2950	4100	innovation rate

The social value accounts for the value created by effective amount of knowledge transfer, not just the part appropriated by innovators. The marginal private value of innovation in equation (18), the future stream of transfer payoff appropriated by the innovator, $q(\alpha)(\beta-\theta)$, enters the value function. In equation (18), the entire payoff from future stream of knowledge transfer, $q(\alpha)(1-\theta)$, enters the social value function.

4 Quantitative Evaluation

In this section, we first calibrate the dynamic model and then quantitatively assess the impact of knowledge appropriability on economic growth.

4.1 Calibration

We calibrate the model to match a set of aggregate and micro-level moments. The micro-level moments include patterns of cross-sectional firm heterogeneity and firm dynamics.[9] Assuming that the economy is on a BGP, we take advantage of the sharp characterization of the properties of the economy along the BGP in Proposition 3 and Lemma 2 to map the parameters to their corresponding moments. The calibration is carried out at annual frequency. One unit of time corresponds to one year in the data. Table 1 displays the calibrated parameter values.[10] Table 2 reports the moments in the model and in the data. The discount rate ρ matches an annual interest rate of 4% and an annual growth rate of 2%.

[9] A more full-blown calibration exercise would require extending the model to include multi-factor production function and richer features such as entry and exit. The key forces of knowledge creation and diffusion would dominate in a richer model.

[10] Taken together, the parameter values ensure that Assumption 1 is satisfied and the firm's problem is well defined.

Table 2: Calibrated moments

Moments	Data	Model	
		Fixed Vol.	Scalable Vol.
Annual interest rate	4%	4%	4%
Annual growth rate	2%	2%	2%
Pareto right tail index	1.06 (2)	2	2
St. dev of growth rates, large firms	0.08	0.08	0.08
Creative destruction probability	2/3	2/3	2/3
Social return/private return	2.6	2.6	2.6

According to the interest rate equation $r = \rho + g$, the discount rate ρ is 0.02.

Standard deviation. Lemma 2 states that the standard deviation of the Brownian motion corresponds to the volatility of growth rates of the largest firms in the economy. Using the Longitudinal Business Database, Davis, Haltiwanger, Jarmin and Miranda (2007) find that the volatility of growth rate for large public-listed firms in the range of 0.05 to 0.01.[11] Therefore, we set the standard deviation $\sigma(\bar{\mu})$ to 0.08. Under the fixed-volatility innovation process $\sigma(\bar{\mu}) = \sigma$, the parameter σ is simply 0.08. Alternatively, under the scalable-volatility innovation process $\sigma(\bar{\mu}) = \sigma\sqrt{\bar{\mu}}$, we back out the parameter σ after obtaining the extent of innovation intensity.

Growth decomposition. The results in Proposition 3 allow us to quantity the amount of growth through innovation and knowledge diffusion. Recall that the productivity distribution in the right tail is tightly linked to the equilibrium contact rate for innovators and the standard deviation of the Brownian motion, $\zeta = \frac{\sqrt{2q(\alpha)}}{\sigma(\bar{\mu})}$. The right tail has a Pareto index of around 1.1 in the U.S. data. We instead target a Pareto right tail index 2. The model-implied contribution of knowledge diffusion to growth is

$$\sigma(\bar{\mu})\sqrt{2q(\alpha)(1-\theta)} = \sigma(\bar{\mu})^2 \zeta = 1.28\%.$$

The residual contribution of innovation to growth is

$$\bar{\mu} = g - \sigma(\bar{\mu})\sqrt{2q(\alpha)(1-\theta)} = 0.72\%.$$

[11]This measure of the volatility of firm growth rate in Davis et al. (2007) excludes short-lived firms and entry and exit. The dispersion in growth rates in the cross section is much higher as it accounts for short-lived firms, as well as entry and exit. The former is the relevant measure for our purpose.

Creative destruction. To calibrate the creative destruction probability θ, we borrow from the recent paper by Garcia-Macia et al. (2019), who estimate the extent of creative destruction of innovation based on the patterns of job creation and job destruction at the firm level. They find that overall, in 1983-1993, 25.5% of aggregate growth is from creative destruction and 13.7% is from new varieties. Decomposing into entrants and incumbents, they find that entrants contribute to 18.6% of growth from creative destruction and 13.7% of growth from new varieties. In the subsequent two decades, they find that the growth from creative destruction increased while the growth from new varieties decreased. From the lens of our model, Considering these patterns, we set the creative destruction probability θ to 2/3, which is in the ballpark of their estimates.[12]

Bargaining weight. The innovator's bargaining weight β shapes the private return of innovation relative to the social return. When the innovators appropriate all gains from knowledge diffusion, the gap between the private return and the social return closes to zero. Bloom et al. (2013) estimate that the social rate of return to R&D is around 55%, while the private return is 21%. In our model, the ratio of social return to private return is

$$\frac{\rho + g - q(\alpha)(\beta - \theta) - \bar{\mu} - \frac{1}{2}\sigma(\bar{\mu})^2}{\rho + g - q(\alpha)(1 - \theta) - \bar{\mu} - \frac{1}{2}\sigma(\bar{\mu})^2} = 2.6.$$

Given that $\rho + g = 4\%$, $q(\alpha)(1-\theta) = 1.28\%$, $\mu = 0.72\%$, $\sigma(\bar{\mu}) = 0.08$, we obtain that $q(\alpha)(\beta - \theta) = -0.4$. Further, given that the calibrated θ is 2/3, the bargaining weight β is 0.3. To gauge whether the calibrated bargaining weight is sensible, we look into the royalty rates in technology licensing deals. Practitioners in licensing transactions commonly apply a "25% rule", which states that the licensees should pay a 25% royalty rate out of the profits from the licensing deals to the licensors. Our calibrated bargaining weight is close to this rule. The model-implied indirect net payoff from knowledge transfer is around -40% in both cases of innovation process.

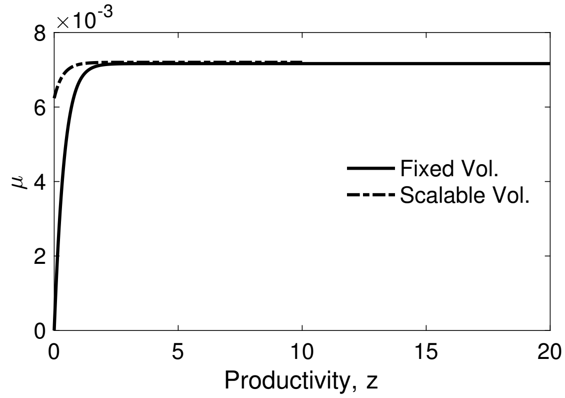
Matching function. We specify a Cobb-Douglas matching function,

$$M(\alpha, 1-\alpha) = \phi\alpha^{\omega}(1-\alpha)^{1-\omega}.$$

The contact rate for innovators is $q(\alpha) = \phi\left(\frac{\alpha}{1-\alpha}\right)^{\omega-1}$ and the contact rate for imitators is $p(\alpha) = \phi\left(\frac{\alpha}{1-\alpha}\right)^{\omega}$. The scale parameter ϕ captures the ease of knowledge diffusion. The

[12] Although we consider a pure endowment economy, our model can be mapped to one in Garcia-Macia et al. (2019). This mapping can be achieved by extending our model to a linear production technology with labor as inputs under monopolistic competition.

Figure 2: Innovation intensity

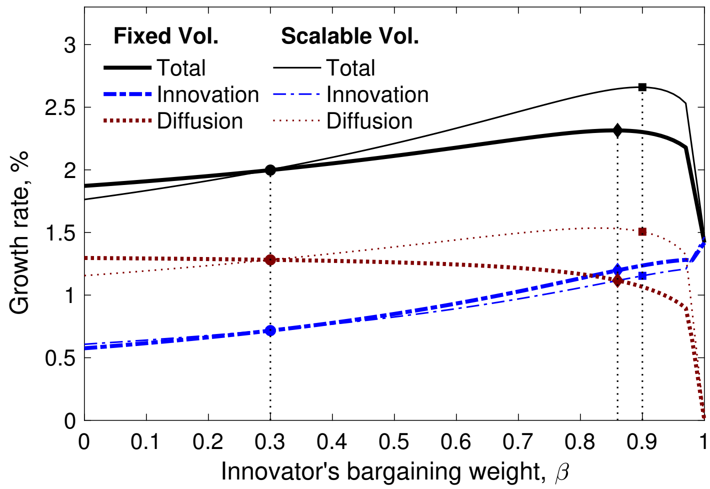


curvature parameter $\omega \in [0,1]$ controls the extent of congestion on the innovator side vis-à-vis the imitator side. If $\omega = 0$, there is no congestion on the imitator side and maximum congestion on the innovator side. Conversely, if $\omega = 1$, there is no congestion on the innovator side but maximum congestion on the imitator side.

Although this type of matching technology is widely examined in studies of labor markets, it is less explored in technology markets. One notable exception is Akcigit, Celik and Greenwood (2016) who look into the patent resale market and use random search to model that market. They also specify a Cobb-Douglas matching function and, using the empirical distribution of the duration until a patent gets sold, find equal amount of congestion on the two sides of the market. Hence, we set ω to 0.5 such that the extent of congestion is equal on the two sides. Further, we carry out sensitivity analysis by varying the congestion parameter later on. The parameter ϕ is calibrated to match the contribution of diffusion to growth. To target a diffusion rate of $q(\alpha)(1-\theta) = (\sigma(\bar{\mu})\zeta)^2/2 = 1.28\%$, in the the case with fixed volatility, we obtain a ϕ of 0.034. The model-implied fraction of firms on the innovating side is roughly 43%. In the case with scalable volatility, we obtain a ϕ of 0.038. The implied fraction of innovating firms is 49%.

Innovation cost function. The innovation function is specified as $c(\mu) = \frac{\gamma}{\varepsilon+1}\mu^{\varepsilon+1}$. We fix the elasticity parameter ε at 1, following the macro and micro estimates of innovation elasticity around unity. We calibrate the innovation cost scale parameter γ to match the contribution of innovation to growth. With a fixed volatility, we obtain a γ of 2950. With a scalable volatility, we obtain a γ of 4100. Figure 2 displays the optimal investment for innovators depending on their distance to the threshold at the calibrated parameters.

Figure 3: Bargaining weight and growth rate

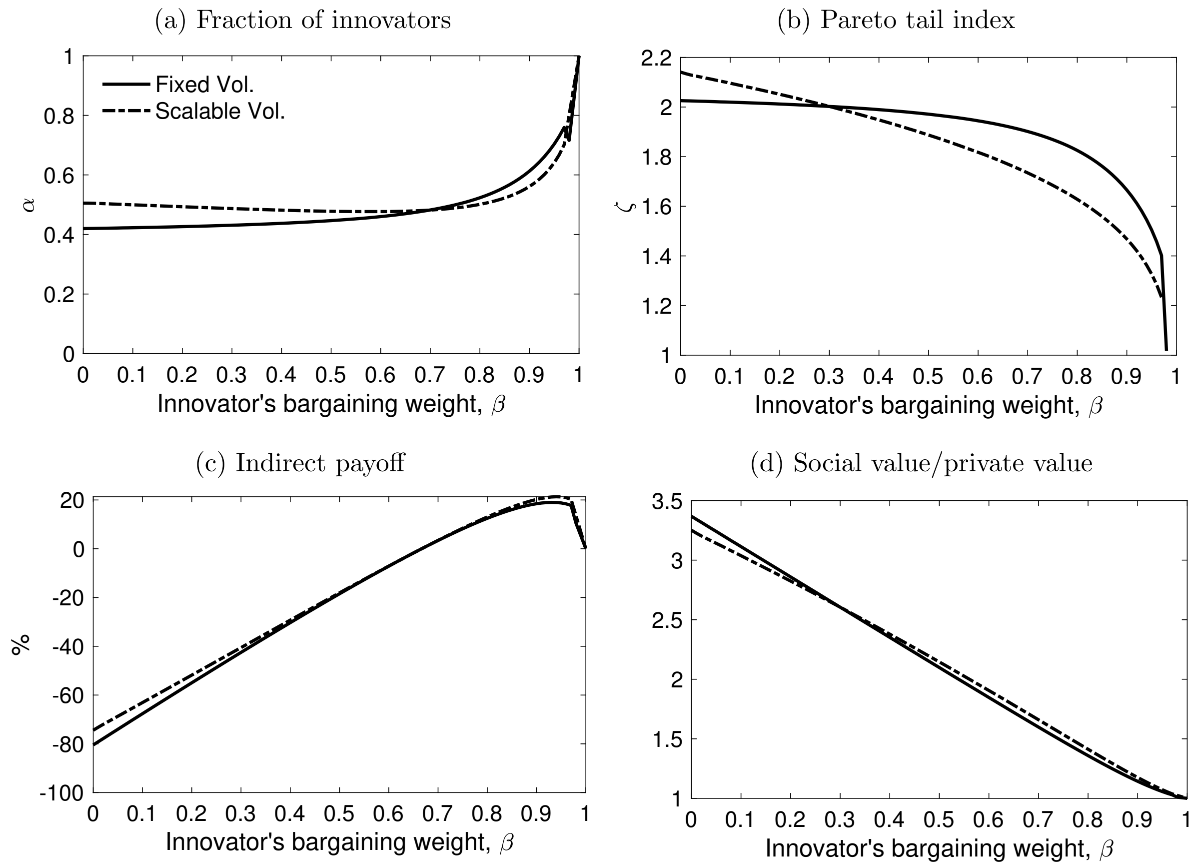


4.2 Counterfactuals: Appropriability and Growth

We now evaluate how the extent of appropriation in knowledge diffusion affects the incentives for innovation and imitation and hence the overall economic growth. We use the calibrated model to carry out the counterfactual exercise of varying the bargaining weights. Figure 3 plots, as the innovator's bargaining weight β increases from 0 to 1, the corresponding growth rate and its decomposition into innovation and diffusion, for both fixed and scalable volatility innovation processes. Figure 4 shows the fraction of innovators, heterogeneity measure, indirect payoff to innovation, and the ratio of social value to private value of innovation as implied by the model.

With fixed volatility innovation process, the growth rate is maximized when the innovator's bargaining weight is 0.86. This corresponds to an effective appropriation by the innovators, $(\beta^* - \theta) / (1 - \theta)$, at 0.58. This is due to, as we increase the innovator's bargaining weight, the diffusion rate responds with a very mild decline while the innovation rate gains by a larger magnitude. On one hand, panel (a) of 4 shows that when we increase the innovator's bargaining weight from 0 to 0.8, the fraction of firms who choose to innovate is very flat. We note that the modeling assumption we have made that imitators are idle in production matters for this result. This assumption makes all imitators identical in value, eradicating the potential distinction between the average imitator and the marginal imitator. In this case, f the marginal innovator's value is not directly affected by its bargaining weight β. The adjustment of the equilibrium fraction of innovators only comes from the imitator's bargaining weight $1 - \beta$. Hence, the elasticity of diffusion to innovator bargaining weight is mild. On the other hand, the elasticity of innovation to innovator bargaining weight is large because the indirect payoff changes quite substantially as we vary the innovator's appropri-

Figure 4: Model-implied moments



ability. Panel (c) shows, as we increase the innovator's organizing weight, the indirect payoff could increase substantially.

With the scalable volatility innovation process, the overall growth rate tends to be maximized by an even higher level of innovator bargaining weight, $\beta^* = 0.9$. The effective appropriation by innovators, $(\beta^* - \theta) / (1 - \theta)$, is 0.7. The response of the innovation rate is similar to the case with fixed volatility. This is because the indirect payoff to innovation is largely driven by the change in appropriability in both cases, as shown in panel (c) of Figure 4. However, the response of knowledge diffusion behaves very differently. As the innovation rate $\bar{\mu}$ increases at a higher innovator appropriation, or in other words as the extent of experimentation increases, the equilibrium level of heterogeneity $\sigma\sqrt{\mu}$ also increases. As shown in panel (b), the Pareto tail index experiences a sharper decrease in the case of scalable volatility compared to the case with fixed volatility. This has a secondary effect on the diffusion rate $\sigma\sqrt{\mu}\sqrt{2q(\alpha)}$. In fact, the magnitude of higher heterogeneity dominates the change in the contact rate, resulting in higher diffusion rate in a large interval of values of bargaining weight.

Figure 5: The effect of congestion and volatility on the optimal bargaining weight

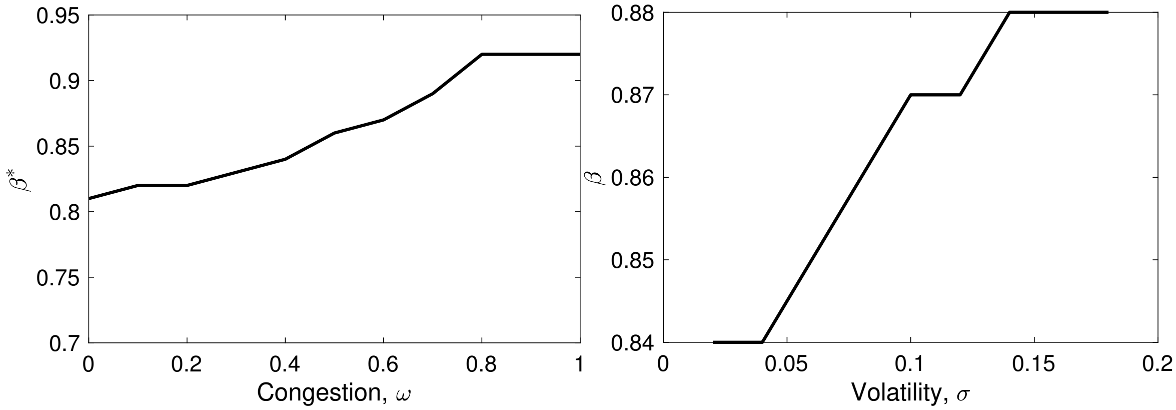
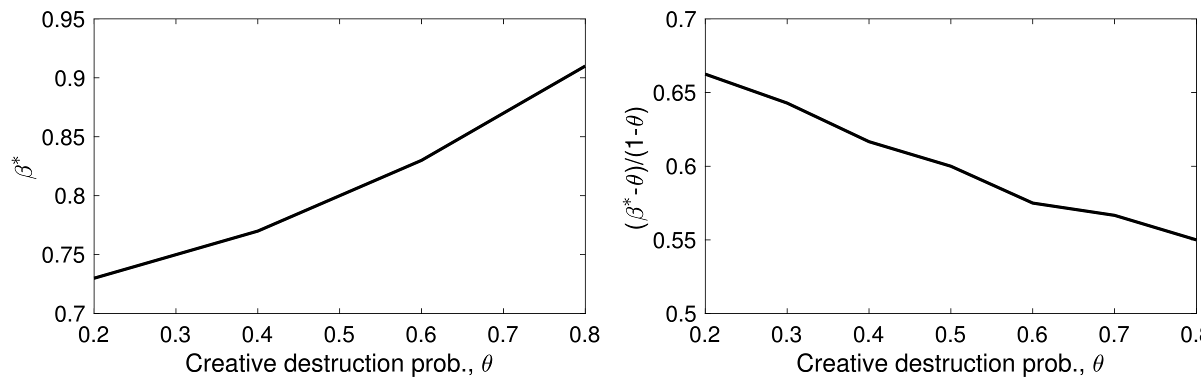


Figure 6: The effect of creative destruction on the optimal bargaining weight

(a) Optimal bargaining weight

(b) Effective optimal bargaining weight



4.3 Discussions

Three parameters are potentially important in determining the growth maximizing degree of innovator appropriation: the extent of learning congestion ω, the innovation outcome dispersion σ, and the creative destruction probability θ. We consider the impact of changing each of the three parameters while keeping other parameters fixed at their calibrated level. Given that the two innovation processes have mild differences, we carry out this sensitivity analysis for the fixed volatility process only.

First, if we were to target a higher level of volatility of innovation process at 0.1 and correspondingly a higher equilibrium level of heterogeneity, it would reenforce that the optimal innovator bargaining weight should be zero. On the contrary, when we reduce the targeted level of volatility of innovation process, for example at a low level 0.01, the optimal innovator bargaining weight would be 0.61. Second, when the weight of innovators in the matching function ω is assumed to be 0, then the optimal innovator bargaining weight is zero. However, if the weight of innovators in the matching function ω is 1, then the optimal innovator bargaining weight is one. This latter case represents the situation where innovators' contact rate is independent of the composition, so entry of marginal firms into this side of the market has no negative external effect. Moreover, as $\omega \to 1$, total matching becomes proportional to the mass of innovators so it is optimal to make it as large as possible. This is accomplished by having larger β. As there are no external effects and higher β internalizes all returns from innovation, the optimal value converges to one. One key insight stands out from this exercise: only when there is a substantial reduction in the volatility of the innovation process, a different conclusion on the growth maximizing innovator bargaining weight might be obtained.

5 Conclusion

To be added.

References

Acemoglu, Daron and Ufuk Akcigit, "Intellectual Property Rights Policy, Competition and Innovation," *Journal of the European Economic Association*, 2012, *10* (1), 1–42.

Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood, "Buy, Keep, or Sell: Economic Growth and the Market for Ideas," *Econometrica*, 2016, *84* (3), 943–984.

Alvarez, Fernando E, Francisco J Buera, and Jr. Lucas Robert E, "Models of Idea Flows," Working Paper 14135, National Bureau of Economic Research June 2008.

Atkeson, Andrew and Ariel Burstein, "Aggregate Implications of Innovation Policy," *Journal of Political Economy*, 2019, *127* (6), 2625–2683.

Benhabib, Jess, Jesse Perla, and Christopher Tonetti, "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier," Working Paper 23095, National Bureau of Economic Research January 2017.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Boldrin, Michele and David Levine, *Against Intellectual Monopoly*, Cambridge University Press, 2008.

Buera, Francisco J. and Ezra Oberfield, "The Global Diffusion of Ideas," *Econometrica*, Forthcoming.

Chari, V. V. and Hugo Hopenhayn, "Vintage Human Capital, Growth, and the Diffusion of New Technology," *Journal of Political Economy*, 1991, *99* (6), 1142–1165.

David, Joel M., "The Aggregate Implications of Mergers and Acquisitions," *Working Paper*, 2017.

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda, "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," *NBER Macroeconomics Annual 2006*, 2007.

Eeckhout, Jan and Boyan Jovanovic, "Knowledge Spillovers and Inequality," *American Economic Review*, 2002, *92* (5), 1290–1307.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow, "How Destructive Is Innovation?," *Econometrica*, 2019, *87* (5), 1507–1541.

Hosios, Arthur J., "On the Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies*, 1990, *57* (2), 279–298.

Jovanovic, Boyan and Glenn M. MacDonald, "Competitive Diffusion," *Journal of Political Economy*, 1994, *102* (1), 24–52.

_ **and Rafael Rob**, "The Growth and Diffusion of Knowledge," *The Review of Economic Studies*, 1989, *56* (4), 569–582.

König, Michael, Jan Lorenz, and Fabrizio Zilibotti, "Innovation vs. imitation and the evolution of productivity distributions," *Theoretical Economics*, 2016, *11* (3), 1053–1102.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, 1997, *65* (6), 1389–1419.

Lucas, Robert E., "Ideas and Growth," *Economica*, 2009, *76* (301), 1–19.

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

_ , "On the Mechanics of Firm Growth," *The Review of Economic Studies*, 2011, *78* (3), 1042–1068.

_ , "Technology diffusion and growth," *Journal of Economic Theory*, 2012, *147* (2), 602 – 622.

_ , "An Assignment Model of Knowledge Diffusion and Income Inequality," 2015.

Michelacci, Claudio, "Low Returns in R&D Due to the Lack of Entrepreneurial Skills," *The Economic Journal*, 2002, *113* (484), 207–225.

Moll, Benjamin and Robert E. Lucas, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Mortensen, Dale T. and Christopher A. Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment," *The Review of Economic Studies*, 1994, *61* (3), 397–415.

Perla, Jesse and Christopher Tonetti, "Equilibrium Imitation and Growth," *Journal of Political Economy*, 2014, *122* (1), 52–76.

_ , _ , and Michael E. Waugh, "Equilibrium Technology Diffusion, Trade, and Growth," Working Paper 20881, National Bureau of Economic Research January 2015.

Shimer, Robert and Lones Smith, "Matching, Search, and Heterogeneity," *The B.E. Journal of Macroeconomics*, 2001, *1* (1), 1–18.

Appendix

A Proofs

A.1 Proof of Lemma 1

First, we divide both sides of the HJB equation (10) for the innovator's value function by $\exp(z)$,

$$r(t)\frac{V^i(z,t)}{\exp(z)} = \max_{\mu}\left\{1 + q(\alpha(t))(\beta-\theta)\left(\frac{V^i(z,t)}{\exp(z)} - \frac{V^c(t)}{\exp(z)}\right)\right.$$
$$\left. -c(\mu) + \mu\frac{V^i_z(z,t)}{\exp(z)} + \frac{1}{2}\sigma(\mu)^2\frac{V^i_{zz}(z,t)}{\exp(z)} + \frac{V^i_t(z,t)}{\exp(z)}\right\}. \tag{21}$$

In equation (21), the flow payoff from knowledge transfer relative to their own productivity, $\frac{V^c(t)}{\exp(z)}$, is decreasing in z. In particular, for firms on the knowledge frontier, this component converges to zero, $\lim_{z\to\infty}\frac{V^c(t)}{\exp(z)} = 0$.

We guess and verify that the innovator's value function is asymptotically affine in $\exp(z)$,

$$\lim_{z\to\infty}\frac{V^i(z,t)}{\exp(z)} = v^i(t). \tag{22}$$

Differentiating equation (22) with respect to z and t, we obtain that

$$\lim_{z\to\infty}\frac{V^i_z(z,t)}{\exp(z)} = \lim_{z\to\infty}\frac{V^i_{zz}(z,t)}{\exp(z)} = v^i(t) \text{ and } \lim_{z\to\infty}\frac{V^i_t(z,t)}{\exp(z)} = v^i_t(t). \tag{23}$$

Plugging the expressions in (22) and (23) into equation (21):

$$r(t)v^i(t) = \max_{\mu}\left\{1 + q(\alpha(t))(\beta-\theta)v^i(t) - c(\mu) + \mu v^i(t) + \frac{1}{2}\sigma(\mu)^2 v^i(t) + v^i_t(t)\right\},$$

which is an ordinary differentiation equation that characterizes $v^i(t)$. The first-order condition with respect to the innovation at the limit is

$$c'(\bar{\mu}(t)) = (1 + \sigma(\bar{\mu}(t))\sigma'(\bar{\mu}(t)))v^i(t).$$

Hence we have verified that $V^i(z,t)$ is asymptotically affine in $\exp(z)$.

We take the limit of the first-order condition (11) for innovation:

$$\lim_{z\to\infty}c'(\mu(z,t)) = \lim_{z\to\infty}(1 + \sigma(\mu(z,t))\sigma'(\mu(z,t)))v^i(t).$$

It must be that $\lim_{z\to\infty} \mu(z,t) = \bar{\mu}(t)$.

A.2 Proof of Proposition 3

On the BGP, given the growth rate g, the interest rate $r(t) = r = \rho + g$. The value functions satisfy $V^i(z+gt,t) = \exp(gt)\, V^i(z)$ and $V^c(t) = \exp(gt)\, V^c$. The innovation threshold also grows at rate g, $\underline{z}(t) = \underline{z} + gt$. Applying these properties, the HJB equations (10) and (12) become

$$\begin{aligned}
\rho V^i(z) =& \max_{\mu} \Big\{ \exp(z) + q(\alpha)(\beta-\theta)\left(V^i(z) - V^c\right) \\
& -c(\mu)\exp(z) + (\mu - g) V^i_z(z) + \frac{1}{2}\sigma(\mu)^2 V^i_{zz}(z) \Big\}, \qquad (24)\\
\rho V^c =& p(\alpha)(1-\beta)\left(\mathbb{E}\left[V^i(\tilde{z})\,|\tilde{z} > \underline{z}\right] - V^c\right). \qquad (25)
\end{aligned}$$

At the innovation threshold $\underline{z}$, the value matching and smooth pasting conditions (7) and (8) become

$$\begin{aligned}
V^c &= V^i(\underline{z}), \\
0 &= V^i_z(\underline{z}).
\end{aligned}$$

Detrending the productivity distribution according to equation (16), we transform the KF equation (13) into an ordinary differential equation, $\forall z \geq \underline{z}$,

$$q(\alpha)(1-\theta) f(z) + g f'(z) - \frac{\partial(\mu(z) f(z))}{\partial z} + \frac{1}{2}\frac{\partial^2\left(\sigma(z)^2 f(z)\right)}{\partial z^2} = 0. \qquad (26)$$

Now at the threshold,

$$q(\alpha)\,\alpha(1-\theta) = g f(\underline{z})$$

The stationary distribution to the right of the threshold does not allow for an analytical solution, due to the varying level of innovation intensity. However, using the result in Lemma 1 that the innovation intensity converges to a constant upper bound in the limit, we can solve the asymptotic distribution in the right tail analytically. When z converges to infinity, equation (26) converges to

$$\lim_{z\to\infty}\left\{ q(\alpha)(1-\theta) f(z) + (g - \bar{\mu}) f'(z) + \frac{1}{2}\sigma(\mu)^2 f''(z)\right\} = 0.$$

The asymptotic stationary distribution in the right tail follows a mixture of Gamma distributions, $\forall z > \underline{z}$,

$$f(z) = \left(B_1 + B_2\left(z - \underline{z}\right)\right)\left(\frac{\exp\left(\underline{z}\right)}{\exp\left(z\right)}\right)^{\zeta},$$

where

$$\zeta = \frac{1}{\sigma\left(\bar{\mu}\right)^2}\left(g - \bar{\mu} \pm \sqrt{\left(g - \bar{\mu}\right)^2 - 2\sigma\left(\bar{\mu}\right)^2 q\left(\alpha\right)\left(1 - \theta\right)}\right). \tag{27}$$

Hence the productivity $Z = \exp(z)$ has a Pareto index ζ in the right tail.

As in Luttmer (2007) and Benhabib et al. (2017), there can potentially exist a continuum of equilibria associated with different growth rates and stationary distributions. When the initial distribution has a bounded support under Assumption 2, we obtain a unique BGP with growth rate.

$$g = \bar{\mu} + \sigma\left(\bar{\mu}\right)\sqrt{2q\left(\alpha\right)\left(1 - \theta\right)}.$$

This is the lower bound of growth rate g derived from equation (27). The Pareto tail index in equation (27) becomes

$$\zeta = \frac{\sqrt{2q\left(\alpha\right)\left(1 - \theta\right)}}{\sigma\left(\bar{\mu}\right)}. \tag{28}$$

To solve for the endogenous innovation decision, we apply again the result in Lemma 1. The value function on the knowledge frontier is asymptotically affine in $\exp(z)$:

$$\lim_{z \to \infty} \frac{V^i\left(z\right)}{\exp\left(z\right)} = v^i, \text{ where } v^i = \frac{1 - c\left(\bar{\mu}\right)}{\rho + g - q\left(\alpha\right)\left(\beta - \theta\right) - \bar{\mu} - \frac{1}{2}\sigma\left(\bar{\mu}\right)^2}. \tag{29}$$

Combing equation (29) with the first-order condition (11) for innovation, we obtain equation (18) for characterizing the innovation upper bound $\bar{\mu}$. Finally, according to the expression in equation (29), Assumption 1 is necessary to ensure that firm's problem is well defined.

A.3 Proof of Lemma 2

Along the BGP, the productivity of innovating firms, relative to the trend, follows the diffusion process:

$$dz - gdt = \left(\mu\left(z\right) - g\right)dt + \sigma\left(\mu\left(z\right)\right)dB.$$

The expected growth rate over a relative short time interval of τ is:

$$\mathbb{E}\left[z_{t+\tau} - z_t - g\tau\right] = \left(\mu\left(z_t\right) - g\right)\tau.$$

For firms on the technology frontier, i.e., $z \to \infty$, their expected growth rate converges to

$$\lim_{z \to \infty} \mathbb{E}\left[z_{t+\tau} - z_t - g\tau\right] = \left(\bar{\mu} - g\right)\tau.$$

The variance of growth rate is

$$\lim_{z \to \infty} \operatorname{var}\left(z_{t+\tau} - z_t - g\tau\right) = \sigma\left(\bar{\mu}\right)^2 \tau.$$

A.4 Proof of Proposition 4

Let $W^i(z,t)$ denote the social value function associated with an innovating firm with productivity z at time t. Let $W^c(t)$ denote the social value function associated with a copying firm at time t. The social value functions satisfy the following HJB equations:

$$\begin{aligned} r(t)\,W^i(z,t) =& \exp(z) + q(\alpha(t))(1-\theta)\left(W^i(z,t) - W^c(t)\right) \\ & - c(\mu)\exp(z) + \mu W^i_z(z,t) + \frac{1}{2}\sigma(\mu)^2 W^i_{zz}(z,t) + W^i_t(z,t), \end{aligned} \tag{30}$$

$$r(t)\,W^c(t) = p(\alpha(t))(1-\theta)\left(\mathbb{E}\left[W^i(z,t)\middle|\, \underline{z} > z(t)\right] - W^c(t)\right) + W^c_t(t). \tag{31}$$

On the BGP, equation (30) simplifies to

$$\rho W^i(z) = \exp(z) + q(\alpha)(1-\theta)\left(W^i(z) - W^c\right) - c(\mu)\exp(z) + (\mu - g)\,W^i_z(z) + \frac{1}{2}\sigma(\mu)^2 W^i_{zz}(z).$$

Taking the limit $z \to \infty$:

$$\lim_{z\to\infty}\frac{W^i(z)}{\exp(z)} = w^i, \text{ where } w^i = \frac{1 - c(\bar{\mu})}{\rho + g - q(\alpha)(1-\theta) - \mu + \frac{1}{2}\sigma(\bar{\mu})^2}.$$

The marginal social value of innovation is then $w^i\left(1 + \sigma(\bar{\mu})\,\sigma'(\bar{\mu})\right)$.