On the Timing of Innovation in Schumpeterian Growth Models*

Gadi Barlevy

Department of Economics
Northwestern University
2003 Sheridan Road
Evanston, IL 60208

e-mail:barlevy@northwestern.edu

January 23, 2003

Abstract

Recently, economists have resurrected the old Schumpeterian notion that recessions can play a useful role in fostering innovation and growth. But in practice, R&D appears to be procyclical. This paper shows that creative destruction, an idea that itself is attributed to Schumpeter, distorts the incentives of entrepreneurs in responding to exogenous fluctuations. As a result, it is possible that equilibrium innovation will be procyclical even though innovation ought to be concentrated during recessions. Thus, what has been previously viewed as an inherently desirable feature of fluctuations in the previous literature – the opportunity to substitute intertemporally – turns out to be a social liability in equilibrium.

*I am grateful to Huw Lloyd-Ellis, Kiminori Matsuyama, Alex Monge, and Marty Eichenbaum for helpful discussions, and to seminar participants at Northwestern, Queens, Fuqua School of Business, NBER, and the Society for Economic Dynamics.

Introduction

In the last decade, economists have resurrected the old Schumpeterian notion that recessions can play a useful role in fostering innovation and growth. Modern reincarnations of this hypothesis are rooted in theories of intertemporal substitution.[1] That is, they assume that productivity-improving activities require resources that could otherwise be employed in production. In downturns, when the return to production is low, there will be a natural incentive to shift resources to productivity-improving activities: the cost of such activities in terms of forgone output or sales is low, while the benefit is high given that economic conditions are expected to improve in the future. This suggests that recessions should serve to facilitate growth, and in some circumstances cyclical fluctuations might even be welfare-improving.

Contrary to this prediction, empirically we observe that certain inputs into technological innovation are in fact concentrated in booms. For example, in surveying the evidence on research and development (R&D), arguably an important source of long-run productivity growth, Griliches (1990) reports that both R&D and its output – patents – are procyclical.[2] Fatas (2000) similarly finds that growth in real R&D expenditures is positively correlated with contemporaneous growth in real GDP, although this correlation is less pronounced in the 1990s. One way to reconcile this finding with the theoretical literature above is to deny the assumptions that give rise to intertemporal substitution in the first place, i.e. to conclude that innovation activities do *not* take resources away from production. For example, Aghion and Saint Paul (1998) show that if productivity-improving activities require produced goods as opposed to factor inputs, innovation will be procyclical. But Aghion and Saint Paul are also quick to dismiss this scenario. Indeed, as noted in Griliches (1984), the main input into R&D is labor, not produced goods. Moreover, the productivity of factor inputs in the goods sector is procyclical — even after correcting for variable utilization as in Burnside, Eichenbaum, and Rebelo (1993) and Basu (1996) — while productivity in the R&D sector (with patents as a proxy for output) is acyclical, according to Griliches (1990). Thus, factor inputs appear to be relatively more productive in the innovation sector during recessions, providing a clear incentive to concentrate

[1] Examples include Hall (1991, 2000), Mortensen and Pissarides (1994), and Gomes, Greenwood, Rebelo (2001) who study job search; Cooper and Haltiwanger (1993), Aghion and Saint Paul (1998), Cooper, Haltiwanger and Power (1999), and Canton and Uhlig (1999) who study technological innovation; and Dellas and Sakellaris (1997), Barlevy and Tsiddon (2000), and DeJong and Ingram (2001) who study human capital accumulation.

[2] The fact that patents are so synchronized with the business cycle might seem somewhat surprising. However, Griliches explains that "the evidence is quite strong that when a firm changes its R&D expenditures, parallel changes occur also in its patent numbers. The relationship is close to contemporaneous with some lag effects which are small and not well estimated (Hall, Griliches, and Hausman, 1986). This is consistent with the observation that patents tend to be taken out relatively early in the life of a research project." (p1674).

innovation in those periods. The purpose of this paper is to understand why innovation might be concentrated in booms despite these incentives.

To be more precise, this paper considers a Schumpeterian growth model in which a fixed amount of resources can be used for either production or innovation. In line with the evidence above, I assume productivity in the goods sector fluctuates over time, while productivity in the innovation sector is constant over time. In this environment, I confirm that it will be optimal to concentrate innovation in recessions, where a recession corresponds to a period of low productivity in the goods sector. This validates the intertemporal substitution view that recessions should facilitate growth. However, I also show that the incentives for agents to shift innovation in response to aggregate fluctuations will be distorted in equilibrium. In particular, I show that if equilibrium profits are highly procyclical, as is the case empirically, innovation in the decentralized economy can be concentrated in booms. The Schumpeterian view is thus correct in arguing that technological considerations dictate that innovation *ought* to be concentrated in recessions, but in equilibrium it might very well be concentrated in booms.

The distortion above is rooted in the fact that I assume growth takes place through creative destruction, an idea that ironically is also credited to Schumpeter. In particular, agents in my model innovate by developing new technologies that use fewer inputs to produce the same amount of output. Once an innovator is successful, he will use his productive advantage to price competitors out of the market, and any previously profitable technology will be rendered obsolete. In judging the value of a successful innovation, then, entrepreneurs take into account only those profits that accrue as a result of their innovation when they are the leading producer. Consequently, entrepreneurs put more weight on profits that accrue in the near future, when they are more likely to be the leading producer, than on profits that accrue further off in the future, when they are likely to be outmoded. As long as shocks are somewhat persistent, the fact that profits are low in recessions and are likely to remain low in the near future can deter firms from undertaking more innovation, even though innovation is relatively less costly during recessions. By contrast, since a social planner does not care which firm accrues profits from a given innovation, he will place relatively more weight on profits that accrue further off in the future, and as such would recognize the full benefits of innovation and undertake more of it in recessions. The inverted timing of innovation is just a special case of the fact that entrepreneurs attach too little value to innovations in recessions and too much value in booms, and are thus likely to engage in less intertemporal substitution than is socially desirable.

The inefficiency above is conceptually different from the inefficiencies that have already been explored in previous work on Schumpeterian growth, e.g. Grossman and Helpman (1991) and

Aghion and Howitt (1992). In particular, previous work has demonstrated that steady-state equilibrium innovation can be too high or too low. This inefficiency concerns the *level* of innovation. But the inefficiency above concerns the *timing* of innovation, independently of its level. That is, even when the average level of innovation is constrained to some suboptimal level, it might still be desirable to concentrate innovation in recessions to achieve growth at a lower cost. The fact that innovation is not sufficiently concentrated in recessions (and might even be concentrated in booms) is thus distinct from the failure to allocate the right amount of resources to innovation on average. To put it another way, an inefficient degree of innovation concerns the *absolute* level of innovation, while an inefficient timing of innovation concerns the *relative* levels of innovation at different points in the cycle.

Note that the inefficient timing of innovation arises because productivity in the goods sector fluctuates over time. As such, we can think of this distortion as a cost of fluctuations, i.e. fluctuations raise the effective cost of undertaking innovation activity. This is a distinct cost of cycles from the one described in Lucas (1987), which is rooted in aversion to risk, as well as from the cost described in Barlevy (2002) in which fluctuations affect long-run growth but where growth is socially efficient. Interestingly, the cost here is due to a feature that previous work has treated as a benefit of fluctuations, namely the ability to concentrate innovation in particular phases of the business cycle. In fact, I show below that it is possible for fluctuations to allow the planner to attain a higher level of welfare than when productivity is stable, but to nevertheless reduce overall welfare in a decentralized equilibrium. Thus, an inherent virtue of cycles is turned into a liability through the private actions of agents.

The paper is organized as follows. Section 1 sets up a benchmark model and characterizes the social optimum and decentralized equilibrium. Section 2 allows for fixed costs of production, which I argue are important in accounting for the high volatility of profits over the cycle. This modification reveals that equilibrium innovation can move in the opposite direction as optimal innovation. Section 3 discusses the welfare implications of the inverted timing of innovation. Section 4 discusses some ways the model can be modified. Section 5 concludes.

1. Basic Setup

I begin with a simple variation on the Grossman and Helpman (1991) model designed to accommodate observed changes in relative productivity between the goods sector and the innovation sector over the business cycle. The model is useful for understanding why private incentives to vary innovation over the cycle differ from those of a social planner. However, because profits in this model exhibit unrealistically low volatility, both optimal and equilibrium innovation are countercyclical. In the next section, I modify the model to account for the high volatility of profits over the business cycle, and show that while optimal innovation remains countercyclical, equilibrium innovation could turn procyclical.

The economy consists of a representative agent, whose instantaneous utility is assumed to be linear in consumption, i.e.

$$U(C_t) = C_t \tag{1.1}$$

By contrast, Grossman and Helpman assume concave utility. The assumption of linear utility is essential for obtaining analytical results. Unfortunately, this assumption also makes it impossible to use the model to realistically assess the quantitative significance of the phenomena I describe. However, as I discuss below, introducing curvature in the utility function is unlikely to overturn my results, so that one could potentially extend the model in a way that is amenable to quantitative analysis. Utility is discounted at rate ρ per unit time.

The agent is endowed with a constant labor endowment L and a constant amount of a fixed factor, e.g. land, per unit time. For convenience, I normalize the supply of the fixed factor to one. To obtain utility from his endowments, the agent must convert them into consumption goods. This transformation occurs in two stages. First, labor must be converted into a series of intermediate goods indexed by $j \in [0, 1]$. Second, intermediate goods, together with the fixed factor, can be converted into a non-storable consumption good.

Turning first to the technology for consumption goods, a producer with F_t units of the fixed factor and x_{jt} units of each intermediate good j can produce final goods according to

$$Y_t = Z_t F_t^{\alpha} X_t^{1-\alpha} \tag{1.2}$$

where Z_t is a measure of productivity in the final goods sector and

$$X_t = \exp\left[\int_0^1 \ln x_{jt} dj\right] \tag{1.3}$$

is a Cobb-Douglas composite of the intermediate goods x_{jt}. Heuristically, think of the intermediate goods as different chemicals, which together combine to make fertilizer. When fertilizer is combined with land, it will yield fruit the agent can consume. The amount of fruit harvested depends on productivity in the final goods sector, Z_t.

To capture the fact that empirically productivity in the goods sector varies systematically over the business cycle, I assume Z_t follows a Markov switching process between two states,

$Z_1 \geq Z_0$, with a constant hazard rate μ. I treat these fluctuations as exogenous. However, one could potentially derive them as a result of some endogenous process.[3]

The role of the fixed factor F above is to allow for diminishing returns to the composite intermediate good X. Diminishing returns plays a similar role in my model as concave utility in the original Grossman and Helpman model. In fact, the model above is equivalent to a model in which there is no fixed factor but the agent has isoleastic utility, just as in Grossman and Helpman. That is, my model is equivalent to a model where

$$U(C_t) = \frac{C_t^{1-\alpha} - 1}{1-\alpha} \tag{1.1$'$}$$

and production of the final good is linear in X_t, i.e.

$$Y = Z_t' X_t \tag{1.2$'$}$$

where $Z_t' = [(1-\alpha) Z_t]^{1/(1-\alpha)}$ is increasing in Z_t for $\alpha \in (0,1)$. Thus, as the model is specified, the cost share of the fixed factor, α, can be interpreted as a coefficient of relative risk aversion. Grossman and Helpman focus on the case of $\alpha = 1$, i.e. log utility. By contrast, the technological restriction that $\alpha \in (0,1)$ in my model implies the implicit coefficient of relative risk aversion in $(1.1')$ is below unity. When I modify the model in the next section to generate greater volatility in equilibrium profits, this equivalence will no longer hold. In that case, I will maintain my original assumptions on preferences and technology, (1.1) and (1.2).

Turning next to the production for intermediate goods, I assume each good $j \in [0,1]$ can be produced from labor according to a linear technology, i.e.

$$x_{jt} = \lambda_{jt} L_{jt} \tag{1.4}$$

where L_{jt} denotes the amount of labor employed in the production of good j at date t. The coefficient λ_{jt} is given by

$$\lambda_{jt} = \lambda^{m_{jt}} \tag{1.5}$$

where $\lambda > 1$ is a constant and m_{jt} is an integer that denotes the generation of technology used in the production of good j at date t. The integer m_{jt} captures the sophistication of production techniques employed in producing good j. The economy begins with an initial technology

[3] For example, following Benhabib and Farmer (1994), one can introduce spillovers across firms that give rise to increasing returns to scale at the aggregate level. In such an economy, there will be equilibria in which the scale of production, and hence the productivity of individual producers, fluctuates endogenously over time.

m_{j0} for each good j. The production of each intermediate good j can always be improved by expending effort towards developing the next generation technology. That is, if R_{jt} units of labor are allocated to research on improving the production of good j, there will be a hazard rate ϕR_{jt} of discovering the next generation $m_j + 1$ in the next instant. This generation will be more productive given $\lambda > 1$. Each time a new generation is discovered, research can begin on the next generation. Let

$$M_t = \int_0^1 m_{jt} dj \tag{1.6}$$

denote the average generation across goods. By the law of large numbers, we surmise $\dot{M}_t = \phi R_t$, where R_t denotes aggregate employment in innovation.

To capture the fact that productivity of innovation is constant over the cycle, I assume ϕ is fixed over time. This last observation is based on patent data, which is probably the best available proxy for output in this sector. Griliches (1990) reports that there is no evidence of systematic changes in the productivity of converting R&D into patents at annual frequencies. As he observes, "Almost all of the systematic short-run variability in aggregate domestic patenting is picked up by fluctuations in the R&D and national defense variables." In particular, changes in real GDP or capacity utilization appear to offer no additional power in explaining the aggregate number of patents. Indeed, since the process of developing new modes of production is so different from the process of producing goods based on an existing blueprint, there is no reason to expect productivity fluctuations in the goods sector to carry over to innovation.

In what follows, it will prove helpful to express the output of final goods directly in terms of labor resources. In particular, suppose each sector uses the same amount of labor, i.e. $L_{jt} = L - R_t$. This symmetric case is natural given the specification for X in (1.3), since both equilibrium and optimal allocations will assign the same number of workers to each intermediate good. Using the fact that the supply of the fixed factor is normalized to 1, output can be expressed directly as a function of labor:

$$Y_t = Z_t X_t^{1-\alpha} = Z_t \left[\lambda^{M_t} (L - R_t)\right]^{1-\alpha} \tag{1.7}$$

The indirect productivity of labor *in terms of final goods* thus depends on both exogenous factors, i.e. Z_t, and endogenous factors, i.e. $\lambda^{(1-\alpha)M_t}$. Assuming a two-stage production structure offers a convenient way to decompose the two terms: productivity in the final goods sector is strictly exogenous, while productivity in the intermediate goods sector is strictly endogenous (and responds to exogenous changes in Z_t). The growth rate of the endogenous component of productivity is given by

$$\frac{d}{dt}\lambda^{(1-\alpha)M_t} = (1-\alpha)\,\dot{M}_t \ln\lambda = (1-\alpha)\,\phi R_t \ln\lambda \tag{1.8}$$

The utility of the agent is finite only if this growth rate does not exceed the discount rate ρ. Thus, a well-defined optimal policy exists only if utility is bounded for any feasible level of innovation, i.e. if

$$\rho > (1-\alpha)\,\phi L \ln\lambda \tag{1.9}$$

As is clear from the above formulation, the inherent tradeoff in this economy is whether to use labor resources to produce output or to promote growth: increasing productivity growth requires a higher level of R_t, which in turn leaves fewer resources to produce goods in the current period. The notion that production and innovation must compete for resources — which ultimately gives rise to this tradeoff — seems reasonable in the context of R&D. In particular, R&D relies heavily on direct labor inputs. For example, Griliches (1984) notes that "R&D is more labor intensive than the average corporate product." While a part of R&D labor expenditures involves scientists and engineers who may not be easily shifted to production, on average about 40% of wage payments in R&D between 1962 and 1998 were allocated to support staff, according to National Science Foundation data. Moreover, engineers who work on applied research and development as opposed to basic research can presumably be shifted more easily to production if necessary, and the bulk of R&D expenditures (roughly 95%) are devoted to the former. Certainly, one can find anecdotal evidence of individual firms that shift resources internally between production and planning in response to changing market conditions.

1.1. The Social Planner's Problem

Intuitively, it seems reasonable that resources in the economy above ought to be shifted into the goods sector when its productivity is high and out of this sector when its productivity is low. Such shifting serves to reduce the cost of productivity growth (in terms of forgone output). This is precisely the idea that underlies the neo-Schumpeterian view. Solving the planning problem that maximizes the utility of the agent confirms that the optimal program in the economy above will indeed follow such a scheme. Formally, let Z_i for $i \in \{0,1\}$ denote the level of productivity at date t, and let M denote the value of the average generation across all goods at this date. The expected utility of the agent as of date t under the optimal path is given by

$$V(Z_i, M) = \max_{R_{t+s}} E_t\left[\int_0^\infty Z_{t+s}\left[\lambda^{M_{t+s}}(L - R_{t+s})\right]^{1-\alpha} e^{-\rho s}ds\right] \tag{1.10}$$

where the maximization problem above is subject to the constraint

$$\dot{M}_{t+s} = \phi R_{t+s}$$

We can rewrite the maximization problem above recursively as follows:

$$\rho V(Z_i, M) = \max_{R \in [0,L]} \left\{ Z_i \left[\lambda^M (L - R) \right]^{1-\alpha} + \mu \left(V(Z_{-i}, M) - V(Z_i, M) \right) + \frac{\partial V}{\partial M} \phi R \right\} \quad (1.11)$$

where Z_{-i} denotes the level of productivity other than Z_i. Given the stationarity of the environment, the planner will choose a constant level of employment R for a given Z, so that the question of choosing an optimal policy reduces to finding a pair of numbers (R_0, R_1). The next proposition establishes the existence of an optimal path and argues it will undertake more innovation when productivity in the final goods sector is low. It is a special case of the more general Proposition 3 below. The proof of that proposition, along with those of all remaining propositions, is contained in an Appendix.

Proposition 1: If (1.9) is satisfied, there exists a unique solution to the social planner's problem, and innovation is (weakly) countercyclical along the optimal path, i.e. $R_0 \geq R_1$.

As noted above, the virtue of shifting innovation to periods when productivity in the goods sector is low is that the economy can achieve growth at a lower cost. At the same time, it should be acknowledged that this benefit is tempered by the fact that varying the amount of resources allocated to production over time is in itself costly. Specifically, diminishing returns to labor implies that varying the amount of labor allocated to production over time will lower average output, due to a Jensen's inequality effect. Alternatively, if we interpret the model in terms of (1.1′) and (1.2′), lowering the cost of growth imposes a cost by forcing a more volatile consumption stream on a risk-averse agent. In the model as specified, this cost is not enough to offset the benefits of intertemporal substitution. But for greater degrees of curvature, this may no longer be true. For example, building on the interpretation of α as a coefficient of risk-aversion in line with (1.1′) and (1.2′), a value of $\alpha > 1$ would imply that innovation ought to be concentrated in periods when productivity in the goods sector is high, i.e. that $R_1 > R_0$ if $Z_1' > Z_0'$. Intuitively, if the agent is very risk-averse, it would be highly desirable to smooth the consumption of the agent. As such, it will be optimal to lower innovation in recessions to offset the fall in productivity and keep consumption high. However, I suspect this result is not really robust. In particular, if we were to introduce an additional margin that can be used to smooth consumption, such as physical capital or storage, the optimal path would likely continue to concentrate innovation when productivity in the final goods sector is low and use the alternative margin to smooth consumption. The reason is that once we separate between production and consumption, there would be a strong incentive to concentrate innovation in recessions in order to maximize production. Formalizing this conjecture, however, requires introducing an additional state variable into the planner's problem. This complication would take us beyond the scope of this paper.

1.2. Decentralized Equilibrium

Next, I examine the decentralized equilibrium in this economy. Production in this economy is carried out by profit-maximizing firms, who pay out profits as dividends to their shareholders. The technology for producing final goods is readily available to any firm. As such, profits in this sector will equal zero in equilibrium. For intermediate goods, firms can take out indefinite patents over production technologies. That is, the firm that discovers the m-th generation for producing good j maintains exclusive rights to this technology. Since no firm would undertake innovation in the absence of patent protection, some monopoly power is essential to sustain growth in equilibrium. Assuming patents last indefinitely is done for mathematical convenience.

An equilibrium for this economy involves a set of prices and quantities that satisfy individual optimization and market clearing. Since the demand for intermediate goods is unit elastic under the Cobb-Douglas aggregator X, any profit-maximizing producer of intermediate goods would wish to charge as high a price as possible. A higher price has no effect on revenue, but lowers the quantity of goods that need to be produced. However, if a producer charges more than the marginal cost of his next most efficient competitor, the latter will underbid him and steal away all of his business. Thus, in equilibrium, only the monopolist with the most productive technology will supply goods, and will charge a price equal to the marginal cost of his most efficient competitor. As Grossman and Helpman observe, incumbent producers benefit less in extending their lead than a new entrant benefits from becoming the new leading producer. Thus, incumbents will not engage in innovation along the equilibrium path, and the next most efficient producer will have the rights to the $(m_{jt} - 1)$-th generation technology.[4] Normalizing the wage to 1, the marginal cost of the next most efficient producer is $\lambda^{-(m_{jt}-1)}$, i.e. the number of labor units he requires to produce a single unit of intermediate good j.

Let p_{jt} denote the price of intermediate good j (relative to the price of labor, which recall was normalized to 1) and $E_{jt} = p_{jt}x_{jt}$ denote total expenditures by final goods producers on good j. Given the Cobb-Douglas specification for X, final goods producers will optimally equalize

[4] As in Grossman and Helpman (1991), this requires that R&D expenditures are not observable, so an incumbent has no incentive ot undertake R&D simply to discourage potential entry.

expenditures across intermediate good j, i.e.

$$E_{jt} = E_t = (1-\alpha) P_t Y_t \tag{1.12}$$

where P_t denotes the price of the final good. With the wage normalized to 1, the cost of production is equal to the number of employed workers, which is just $\lambda^{-m_{jt}} x_{jt}$. Since $x_{jt} = E_j / p_{jt}$, this cost is equal to $\lambda^{-1} E_t$. Hence, the profits of the incumbent firm that supplies good j are just

$$\pi_{jt} = \left(1 - \lambda^{-1}\right) E_t = \left(1 - \lambda^{-1}\right)(1-\alpha) P_t Y_t \tag{1.13}$$

Note that profits are the same for all intermediate goods, regardless of m_{jt}. Using the fact that total spending on consumption goods must equal the income of the representative agent in equilibrium, we can express $P_t Y_t$ as the sum of aggregate profits Π_t and payments to factors, i.e.

$$\begin{aligned} P_t Y_t &= \Pi_t + p_{Ft} F + L \\ &= \int_0^1 \pi_{jt} dj - R_t + p_{Ft} F + L \end{aligned}$$

where p_{Ft} denotes the price of the fixed factor at date t. Substituting in for the expression for π_{jt} above and using the fact that $p_{Ft} F = \alpha P_t Y_t$ allows us to express individual profits π_{jt} strictly in terms of employment in innovation R_t:

$$\pi_{jt} = (\lambda - 1)(L - R_t) \tag{1.14}$$

Note that nominal profits do not depend on Z_t. Thus, if we fixed the level of innovation R, profits would be just as cyclically volatile as the numeraire good, labor. This implies that the volatility of profits should be commensurate with the volatility of the *cost* of R&D together with the volatility of R&D *activity* R. In practice, though, profits are far more volatile than implied by these two series. This observation will figure prominently below.

In considering whether to undertake innovation, entrepreneurs anticipate that if they succeed in innovation, they will earn profits (1.14) as long as their technology is the most advanced. Define $\mathbb{I}_{jt}$ as an indicator variable that equals 1 if the entrepreneur is the leading-edge producer of good j and zero otherwise. To ascertain the value of a successful innovation, I follow Lucas (1978) in pricing an asset. That is, suppose there was a market for claims on the profits of a firm that successfully innovated the production process of good j, and denote the price of this claim at date t by v_{jt}. Since the representative agent will own all claims in equilibrium, v_{jt}

must leave the agent indifferent between buying a marginal claim and selling a marginal claim given he already owns all claims. This indifference condition yields the following

$$\begin{aligned} v_{jt} &= E_t\left[\int_0^\infty \mathbb{I}_{j,t+s} \cdot \frac{U'\left(C_{t+s}\right)/P_{t+s}}{U'(C_t)/P_t}\pi_{t+s}e^{-\rho s}ds\right] \\ &= E_t\left[\int_0^\infty \mathbb{I}_{j,t+s} \cdot \frac{P_t}{P_{t+s}}\pi_{t+s}e^{-\rho s}ds\right] \end{aligned} \tag{1.15}$$

The expectation above is taken over all possible paths for Z_t and all realizations of when entrants succeed in innovating good j. At each instant, non-incumbent firms choose R_{jt} to maximize the expected value from a successful innovation net of R&D costs, i.e. $\phi R_{jt}v_{jt} - R_{jt}$. Since resources are finite and this expression is linear in R_{jt}, it follows that $\phi v_{jt} \leq 1$ in equilibrium, with strict equality at any date t for which $R_{jt} > 0$. I will henceforth limit attention to symmetric equilibria in which R_{jt} is the same across all goods j, i.e. $R_{jt} = R_t$. Since profits π_t are the same for all intermediate goods, this ensures the value of a successful innovation v_{jt} will be the same for all j, i.e. $v_{jt} = v_t$. In addition, I limit attention to Markov-perfect equilibria, i.e. equilibria in which the level of aggregate research R_t depends only on the current level of productivity Z_t. Such an equilibrium can be summarized by a pair of numbers (R_0, R_1). Since the optimal path maintains a constant level of innovation for a given level of productivity, it seems natural to focus on such equilibria.

Given that π_t can be expressed as a function of R_t, we should be able to express the value of a successful innovation v_t above strictly in terms of (R_0, R_1). To do this, we need to first express the price of final goods P_t in terms of R_t. Since the production of final goods is competitive, the price P equals the minimum cost to produce a single unit of the good in equilibrium, i.e.

$$P_t = \min_{x_{jt}, F_t}\left\{\int_0^1 p_{jt}x_{jt}dj + p_{Ft}F_t\right\} \tag{1.16}$$

s.t.

$$Z_t F_t^\alpha \left(\exp\left[\int_0^1 \ln x_{jt}dj\right]\right)^{1-\alpha} = 1$$

Using the fact that $p_{jt} = \lambda^{-(m_{jt}-1)}$ and the fact that $p_{Ft} = \alpha P_t Y_t$ in equilibrium, a little algebra yields

$$P_t = \frac{\lambda\left(L - R_t\right)^\alpha}{\left(1-\alpha\right)Z_t\lambda^{(1-\alpha)M_t}} \tag{1.17}$$

The value of a successful innovation v_t can thus be expressed as the following integral, where for convenience I suppress all references to j:

$$\begin{aligned} v_t &= E_t\left[\int_0^\infty \mathbb{I}_{t+s}\cdot\frac{Z_{t+s}}{Z_t}\left(\frac{L-R_t}{L-R_{t+s}}\right)^\alpha\left(\frac{\lambda^{M_{t+s}}}{\lambda^{M_t}}\right)^{1-\alpha}(\lambda-1)(L-R_{t+s})\,e^{-\rho s}ds\right] \\ &= \text{constant}\cdot E_t\left[\int_0^\infty \mathbb{I}_{t+s}\cdot Z_{t+s}\left[\lambda^{M_{t+s}}(L-R_{t+s})\right]^{1-\alpha}e^{-\rho s}ds\right] \end{aligned} \tag{1.18}$$

Note the similarly between (1.18) and (1.10). Entrepreneurs in the decentralized environment value a successful innovation as a discounted sum of future output, just as the social planner. The key difference, other than the scaling constant, is that an entrepreneur only values output at date t if he is the leading edge producer, as reflected by the indicator variable $\mathbb{I}_{t+s}$. As a result, output that is produced further in the future is more heavily discounted under (1.18) than under (1.10), since $E_t\left[\Pr\left(\mathbb{I}_{t+s}=1\right)\right]$ is decreasing with s.

The last observation helps to explain why the decentralized economy might fail to optimally substitute over the cycle. Suppose it were optimal to maintain a countercyclical path for innovation, i.e. to set $R_0 > R_1$. Then output in recessions would be lower than in booms, i.e. $Z_0(L-R_0)^{1-\alpha} < Z_1(L-R_1)^{1-\alpha}$. Since entrepreneurs attach more weight to current economic conditions, it follows that they will attach too little value to successful innovations during recessions and too much value to successful innovations during booms. Hence, starting from the optimal path, entrepreneurs will have incentive to shift some innovation to periods of high productivity, and the relative levels of R_0 and R_1 in equilibrium will be suboptimal. Note that the inefficiency concerns the relative levels of innovation at different stages of the cycle, not their absolute levels. As such, the inefficiency above is conceptually distinct from the observation in previous work that there might be too much or too little innovation in equilibrium in Schumpeterian growth models, which is a statement about the absolute level of innovation.

To solve for the equilibrium path for innovation, I will need to obtain an analytical expression for v_t. Towards this end, note that for any function $X(Z_t)$, the value of the integral

$$W(Z_i,M) = E_t\left[\int_0^\infty \mathbb{I}_{t+s}\cdot\lambda^{(1-\alpha)M_{t+s}}X(Z_{t+s})\,e^{-\rho s}ds\right]$$

subject to $\dot{M}_{t+s}=\phi R_{t+s}$ can be characterized by the following recursive equation:

$$\begin{aligned} \rho W(Z_i,M) &= \lambda^{(1-\alpha)M}X(Z_i)+\mu\left(W(Z_{-i},M)-W(Z_i,M)\right) \\ &\quad +\frac{\partial W}{\partial M}\phi R_i - W(Z_i,M)\,\phi R_i \end{aligned} \tag{1.19}$$

The method of undetermined coefficients confirms that $W(Z_i,M)=w_i\lambda^{(1-\alpha)M}$ where

$$w_i = \frac{\omega(R_{-i})X(Z_i)+\mu X(Z_{-i})}{\omega(R_i)\,\omega(R_{-i})-\mu^2}$$

for

$$\omega(R) = \rho + \mu + (1 - (1-\alpha)\ln\lambda)\,\phi R$$

The value of a successful innovation can therefore be expressed solely in terms of R_0 and R_1. In particular, if the current level of productivity is equal to Z_i, then the value of a successful innovation is given by

$$v_i(R_i, R_{-i}) = (\lambda - 1)\frac{\omega(R_{-i})(L - R_i) + \mu\frac{Z_{-i}}{Z_i}(L - R_{-i})^{1-\alpha}(L - R_i)^{\alpha}}{\omega(R_i)\,\omega(R_{-i}) - \mu^2}\lambda^{(1-\alpha)M_t} \tag{1.20}$$

Solving for a symmetric Markov-perfect equilibrium amounts to looking for pairs $(R_0, R_1) \in [0, L] \times [0, L]$ such that $\phi v_i(R_i, R_{-i}) \leq 1$, with strict equality if $R_i > 0$. The next proposition provides a condition for the existence and uniqueness of a symmetric Markov equilibrium, and shows that despite the fact that private agents undervalue innovations during recessions, the equilibrium path for innovation remains countercyclical.[5]

Proposition 2: Innovation along the equilibrium path is weakly countercyclical in any symmetric Markov-perfect equilibrium, i.e. $R_0 \geq R_1$ along any equilibrium path. Moreover, if $\lambda < e^{\frac{1}{1-\alpha}}$, there exists a unique symmetric Markov-perfect equilibrium.

The reason that equilibrium innovation remains countercyclical is connected to the fact that profits along the equilibrium path of this economy are not sufficiently volatile. Recall from above that holding the level of innovation R constant over the cycle implies profits would be as volatile as the cost of innovation. Thus, other things equal, both profits and the cost of R&D decline at the same rate during recessions. However, in contemplating whether to undertake more innovation, entrepreneurs compare current costs not to current profits but to a discounted sum of future expected profits. While profits in a recession fall to the same extent as the cost of innovation, expected future profits will not fall to the same extent as long as the process Z_t is stationary. Consequently, the discounted value of an innovation will fall by less than the cost of innovation, and agents have incentive to undertake more innovation in recessions. Of course, this does not imply that the equilibrium path for innovation will coincide with the optimal path. To the contrary, the intuition above suggests they are probably different. But since neither the equilibrium nor the optimal path can be solved in closed form, it is not possible to compare

[5] Canton and Uhlig (1999) also argue that equilibrium innovation will be countercyclical in a related model. Aghion and Saint Paul (1998) also claim that equilibrium innovation is countercyclical when innovation activity requires sacrificing production. But they assume the same firm undertakes all innovation, so there is nothing in their model that distorts the incentives for intertemporal substitution that could lead to procyclical innovation.

them analytically. Numerical simulations for various parameter values suggest that equilibrium innovation is systematically less volatile than optimal innovation. A similar pattern can be seen in the numerical simulations reported by Canton and Uhlig (1999).

In sum, the model reveals that the incentives of agents to substitute intertemporally will be distorted, but given that profits are not very volatile in the model, this distortion is not sufficiently pronounced to reverse the direction of innovation over the cycle. Empirically, of course, profits are far more volatile than the cost of innovation. A recent discussion of the highly procyclical nature of profits is contained in Rotemberg and Woodford (1999), although this pattern has also been widely documented in previous work. By contrast, the cost of R&D is far less volatile over the business cycle. For example, Mansfield (1987) constructs R&D price indices for the period covering 1969-1983. He finds that at high frequencies, the R&D deflator is closely synchronized with the implicit GDP deflator, implying real R&D costs are not very cyclically sensitive. This is not surprising given that nearly half of all R&D expenditures go to labor, and it is well-known that real wages are only mildly procyclical. In the next section, I discuss how to modify the model so that it is more consistent with empirical evidence on the volatility of profits over the cycle. As might be expected from the discussion above, more volatile profits will allow equilibrium innovation to be procyclical.

Before turning to this modification, though, I digress to remark on the uniqueness of the equilibrium in Proposition 2. Note that I need to impose a restriction on λ to ensure the equilibrium is unique. To see why multiple equilibria might arise, suppose we start at some equilibrium and conjecture that entrepreneurs decide to undertake more innovation for all goods j. On the one hand, more rapid innovation of other goods increases the value of a successful innovation of any particular intermediate good, since the marginal product of each good rises when there are more of the other intermediate goods available. This validates the decision to undertake more innovation on each individual good, so that an equilibrium with a higher level of innovation might be sustainable.[6] However, more innovation in one's own sector reduces the expected duration of being the leading-edge producer. To rule out multiple symmetric equilibria, the latter effect must dominate. This requires that successful innovation does not improve productivity in remaining sectors too dramatically, i.e. that λ is not too large.[7]

[6] Shleifer (1986) and Francois and Lloyd-Ellis (2002) exploit this complementarity to study animal spirits a closely related model.

[7] Note that as α tends to 1, the bound on λ in Proposition 2 tends to ∞. Thus, multiple equilibria can be ruled out when the coefficient of relative risk aversion is greater than or equal to 1. This explains why Grossman and Helpman (1991), who consider log utility, obtain a unique equilibrium in their model.

2. Fixed Costs and the Volatility of Profits

The previous section established that the creative destruction aspect of Schumpeterian growth models will cause entrepreneurs to undervalue innovations in recessions and overvalue innovations in booms. Despite this, the equilibrium path remains countercyclical. Since this is due to the fact that profits in equilibrium are counterfactually low, we would want to modify the model in a way that would allow for more volatile profits along the equilibrium path. There are various ways to potentially make profits more volatile in the model. However, one of the arguably more important considerations that accounts for the high volatility of profits in practice is fixed costs of production. The reasoning for this is laid out in Ramey (1991). She defines the *profit rate* as the ratio of profits to sales, and shows that this ratio is procyclical in the data. However, in the absence of fixed costs, this ratio is equal to the price-cost margin, i.e.

$$\frac{\pi_t}{E_t} = 1 - \frac{1}{\lambda}$$

Since the markup λ is constant in my model, profit rates will be as well. Empirically, most estimates of the markup suggest it is either acyclical, e.g. Ramey (1991), or countercyclical, e.g. Rotemberg and Woodford (1999). Thus, any feature that generates more volatile profits by increasing the volatility of sales would be unable to explain the procyclicality of profit rates. Instead, as Ramey (1991) concludes, "the only way to reconcile countercyclical markups with procyclical profit rates is to allow for the presence of significant fixed costs" (p137). To see this, suppose we were to introduce a fixed cost FC that does not vary with the amount of sales. The profit ratio in the model would then equal

$$\frac{\pi_t}{E_t} = \left(1 - \frac{1}{\lambda}\right) - \frac{FC}{E_t} \tag{2.1}$$

which increases with E_t. Evidently, subtracting a constant amount from profits makes profits more volatile, but since it has no effect on the volatility of sales, profit rates will appear more procyclical. Rotemberg and Woodford (1999) similarly cite overhead costs as a way to generate procyclical profits despite countercyclical markups. Other evidence that confirms the quantitative importance of fixed costs includes Basu (1996), who finds that production functions are non-homothetic in a way that suggests the presence of overhead costs.

Suppose then that for any intermediate good j, an entrepreneur must incur a fixed cost K_t per unit time in order to commence production. The fact that the fixed cost must not vary with sales requires that the fixed cost be denominated in units of output. Thus, I will assume the setup cost involves K_t units of the final good per unit time. Since the economy grows over time, I will also need to assume that the fixed cost grows at the same rate as the economy so that we do not eventually outgrow it. Thus, let

$$K_t = \lambda^{(1-\alpha)M_t}\kappa \tag{2.2}$$

for some constant κ. One can motivate the growth of the fixed cost by assuming that entrepreneurs need more space to handle the bigger sales volume as the economy becomes more productive. To ensure that equilibrium profits are positive, we must also make sure that fixed costs are not so large that they exceed either the revenue of producers or the total amount of final goods produced.

To see the effects of fixed costs on the optimal path for innovation, note that the planner's problem (1.11) in the presence of fixed costs can now be written as

$$\rho V(Z_i, M) = \max_{R\in[0,L]} \left\{ \begin{array}{r} Z_i \left[\lambda^M (L-R)\right]^{1-\alpha} - \lambda^{(1-\alpha)M}\kappa + \\ \mu \left(V(Z_{-i}, M) - V(Z_i, M)\right) + \dfrac{\partial V}{\partial M}\phi R \end{array} \right\} \tag{2.3}$$

Since the fixed cost grows at the same rate as the economy, the planner takes into account the fact that faster innovation will increase the fixed cost of production. Thus, netting out the fixed cost will lead the planner to choose a different path from the one that solves (1.11). However, since all intermediate goods are produced at every point in time, the fixed cost is incurred regardless of the level of productivity Z_t. Hence, while the fixed cost could affect the optimal level of innovation, there is no obvious reason that it should affect the optimal timing of innovation. In the next proposition, I confirm that the optimal path remains countercyclical for an interior optimum (i.e. where $R_t > 0$ for all t). For the more general case, I can only establish this claim for small values of κ, although I suspect that the statement of the proposition is true for large values of κ as well.

Proposition 3: If (1.9) is satisfied, there exists a unique solution to the social planner's problem. If the optimal path for innovation involves positive levels of innovation at all dates, innovation will necessarily be countercyclical along the optimal path, i.e. $R_0 > R_1$. If κ is sufficiently small, the optimal path will be weakly countercyclical along the optimal path even if the optimal path involves periods of zero innovation, i.e. $R_0 \geq R_1$.

Next, I turn to the decentralized economy. Once again, I begin with the pricing decisions of the incumbent firm. As before, the firm will try to increase its price as much as possible. Rather than setting its price to the marginal cost of its next most efficient competitor, however, it will now set its price to the level at which its competitor, who also faces a fixed cost, breaks even. Thus, an entrepreneur with productivity λ^m will choose a price p_t that solves

$$\left(1 - \frac{1}{\lambda^{m-1} p_t}\right) E_t = P_t K_t \tag{2.4}$$

Solving for p_t and substituting into the profits of the leading producer yields

$$\pi_{jt} = \frac{\lambda - 1}{\lambda} \left(E_t - P_t K_t\right) \tag{2.5}$$

which again is the same for all j. Using the fact that the representative agent spends all of his income on final goods, we can express profits π_{jt} in terms of R_t analogously to (1.14):

$$\pi_{jt} = (\lambda - 1)\left(L - R_t - P_t K_t\right) \tag{2.6}$$

The price of final goods P_t can once again be obtained by solving for the minimum cost of producing a single unit of the final good. A little algebra shows that

$$P_t = \frac{\lambda}{\lambda^{(1-\alpha)M_t}} \frac{L - R_t}{(1-\alpha) Z_t (L - R_t)^{1-\alpha} - \kappa} \tag{2.7}$$

Note that for $\kappa = 0$, this expression is equivalent to (1.17). Substituting for P_t yields the following expression for profits:

$$\pi_{jt} = (\lambda - 1)\left(1 - \frac{(\lambda + 1)\kappa}{(1-\alpha) Z_t (L - R_t)^{1-\alpha} - \kappa}\right)(L - R_t) \tag{2.8}$$

In contrast with the previous section, nominal profits now do depend on Z_t. In particular, for a fixed level of innovation R, profits rise with Z_t. Thus, other things equal, a recession will be associated with a disproportionate fall in profits relative to the cost of R&D. Moreover, the larger is κ, the more responsive profits will be to changes in aggregate productivity Z_t, i.e. $\frac{d\pi_t}{dZ_t}$ is increasing in κ. For sufficiently large fixed costs, then, we would anticipate that profits fall enough during recessions so that the present discounted value of profits v_t will fall relative to the cost of R&D, inducing agents to undertake *less* rather than more innovation in recessions.[8]

Formally, we can use (1.19) to derive $v_i(R_i, R_{-i})$ and search for an equilibrium pair (R_0, R_1) in $[0, L] \times [0, L]$ for which $\phi v_i(R_i, R_{-i}) \leq 1$, with equality if $R_i > 0$. Unfortunately, the expression for $v_i(R_i, R_{-i})$ is too unwieldy and does not allow me to characterize the equilibrium

[8] In contrast to the previous section where the markup was constant over the cycle, now the markup varies with Z_t. More specifically, the equilibrium markup will be countercyclical, which accords with evidence in Rotemberg and Woodford (1999). Intutively, since profits are higher in booms, a given fixed cost is less effective in discouraging entery by rivals, which forces the leading producer has to lower his price – and thus his markup – to deter entry. This explanation is distinct from those Rotemberg and Woodford offer to explain the countercyclicality of markups, but is similar in implying that the threat of competition becomes more significant during booms.

analytically. Below, I modify the model in a way that allows me to establish a formal link between the size of the fixed cost and the potential existence of an equilibrium with procyclical innovation. However, I first provide a numerical example that establishes similar results in the model as specified. The parameter values are displayed in Table 1 below. As noted earlier, the fact that I assume linear utility (implying an infinite elasticity of intertemporal substitution) makes it difficult to match the model to actual data. Still, to the extent possible, I chose reasonable values for various parameters, such as the discount rate ρ, the average duration of a cycle as reflected in μ, the markup λ, and the volatility of Z_t.[9]

Table 1

ρ	0.04	α	0.33
μ	0.20	Z_0	0.95
ϕ	0.01	Z_1	1.05
λ	1.09	L	500

For the parameters in Table 1, the unique symmetric Markov equilibrium when $\kappa = 0$ is given by $(R_0, R_1) = (41.7, 37.9)$. In line with Proposition 2, equilibrium innovation is countercyclical. As we raise κ, the symmetric Markov equilibrium remains unique. In addition, raising κ lowers the level of innovation, not surprisingly since the fixed costs cuts into profits and thus lowers the value of a successful innovation. But, more importantly, for large enough fixed costs, the timing of innovation in equilibrium will be reversed. To be more precise, there exists a cutoff $\kappa^* \approx 18.86$ at which $R_0 = R_1$. For any value of κ below this cutoff, equilibrium innovation is countercyclical, and for any value of κ above this cutoff for which an interior equilibrium exists, equilibrium innovation is procyclical. Thus, for example, when $\kappa = 19$, the equilibrium path for innovation is given by $(R_0, R_1) = (0.4, 0.5)$. Now, suppose we were to shift to the alternative path $(R_0', R_1') = (0.5, 0.4)$, i.e. we reverse the timing of innovation. This leaves the unconditional growth rate for this economy unaffected, and thus cannot ameliorate any inefficiency that is due to a suboptimally high or suboptimally low level of innovation. However, if we calculate the unconditional expected utility of the agent under the two paths, we find that the agent is better off with the countercyclical path. This example underscores the fact that the inefficient timing of innovation does not reflect an inefficiency in the absolute level of innovation, but in the relative values of innovation at different points in time.

[9] Note that these parameters violate (1.9), so there is no well-defined optimal policy, although there is a well-defined equilibrium. For higher values of ρ, one can construct examples of procyclical equilibria in which (1.9) is satisfied. This reveals a limitation of linear utility. In particular, concave utility serves to raise the effective discount rate, since marginal output in the future is less valued given diminishing marginal utility and growth.

To establish a formal link between the size of the fixed cost and the possibility of procyclical equilibrium innovation, I now turn to a modification of the model that yields a more convenient expression for $v_i(R_i, R_{-i})$. The main difficulty with the expression for P above involves the intercept term κ in the denominator. To get around this, suppose that the production structure is such that only the incumbent has to incur the fixed cost K_t, i.e. any producer who wishes to steal away customers from the leading edge producer does not have to incur the fixed cost. This variation is plausible if the fixed cost reflects a cost of acquiring information. For example, suppose the producer has to learn the exact specifications that final goods producers require to produce at a given point in time. Rival producers can then just observe what goods the leading producer is selling and avoid having to do the research themselves. That said, the real reason for this assumption is not realism – in practice a significant fraction of fixed costs has nothing to do with information acquisition – but to obtain a more convenient functional form for $v_i(R_i, R_{-i})$ with similar qualitative properties. If only the incumbent incurs the fixed cost K_t, he will once again charge a price equal to the marginal cost of his next most efficient producer, rather than the price that satisfies (2.4). Solving for profits and the price of final goods yields the following expression for the value of a successful innovation:

$$v_i(R_i, R_{-i}) = (\lambda - 1)\frac{\omega(R_{-i})(L - R_i) + \mu\frac{Z_{-i}}{Z_i}(L - R_{-i})^{1-\alpha}(L - R_i)^{\alpha}}{\omega(R_i)\,\omega(R_{-i}) - \mu^2} - \frac{\omega(R_{-i}) + \mu}{\omega(R_i)\,\omega(R_{-i}) - \mu^2}\frac{\lambda^2 (L - R_i)^{\alpha}\kappa}{(1-\alpha) Z_i} \tag{2.9}$$

The only way in which the expression above differs from the case in which all firms incur the fixed cost involves the second term above, i.e. the present discounted value of the fixed cost. However, the derivative of this expression with respect to κ and Z has the same sign in both cases. Henceforth, I will let $v_i(R_i, R_{-i})$ denote the expression in (2.9). I begin with the following lemma:

Lemma: Suppose $\lambda < e^{\frac{1}{1-\alpha}}$. Then for any $\kappa > 0$, there exists a unique $R^* < L$ such that $\phi v_0(R^*, R^*) = \phi v_1(R^*, R^*)$. Moreover, there exists a $\kappa^* > 0$ such that $\phi v_i(R^*, R^*) < 1$ for $\kappa < \kappa^*$ and $\phi v_i(R^*, R^*) > 1$ for $\kappa > \kappa^*$.

The lemma above establishes that for any fixed cost κ, there is a unique level of innovation R^* that leaves the value of a successful innovation v constant over time when innovation is held constant. From the proof of the lemma, one can show that this value $v(R^*, R^*)$ increases with κ, ranging from 0 to infinity, so that there must be some κ^* for which it equals $1/\phi$. As the next proposition establishes, if κ is greater than κ^*, we are assured of finding a pair (R_0, R_1) where $R_1 > R_0$ and which satisfies the condition that $\phi v_i(R_i, R_{-i}) = 1$ for both $i \in \{0, 1\}$.

Proposition 4: Suppose $\lambda < e^{\frac{1}{1-\alpha}}$. If $\kappa > \kappa^*$, where κ^* is defined in Lemma 2 , there exists a pair (R_0, R_1) where $R_1 > R_0$ such that

$$\phi v_i (R_i, R_{-i}) = 1$$

Proposition 4 comes close to establishing the conjecture above, namely that for sufficiently large fixed costs, equilibrium innovation will be procyclical. In particular, it assures us that for sufficiently large fixed costs, we can always find a solution for the system of equations associated with an interior equilibrium in which $R_1 > R_0$. However, this implies neither the existence nor the uniqueness of an equilibrium in which innovation is procyclical. In terms of existence, the proof of Proposition 4 does not guarantee that the solution (R_0, R_1) is positive. However, at least numerically, it appears as though assigning a large enough value for L is enough to insure an interior solution, mirroring a similar result in Grossman and Helpman (1991) concerning the existence of an interior equilibrium in the case of no fixed cost (and constant productivity Z_t). In terms of uniqueness, I can no longer establish an analogous result to Proposition 2 above. In fact, one can show that the set $\{(R_0, R_1) \mid \phi v_i (R_i, R_{-i}) = 1\}$ can have multiple solutions beyond the one that is captured in Proposition 4. However, these additional solutions appear to be pathological and do not correspond to actual equilibria. In particular, they appear to involve very high levels of innovation for which the revenue of intermediate goods producers does not cover their fixed costs. In other words, profits are negative in these candidate equilibria, which is not an equilibrium given the option for entrepreneurs to shut down. For all of the parameter values I experimented with in which there was an interior equilibrium, it was in fact unique, and the equilibrium appeared to change continuously with κ from one with countercyclical innovation to one with procyclical innovation.

To conclude, introducing fixed costs of production, which causes profits to be more volatile, reveals that equilibrium innovation can be procyclical even when innovation should optimally be concentrated during booms. This helps to reconcile the observation that innovation is procyclical with evidence on productivity in both the goods sector and the innovation sector that suggest innovation ought to be countercyclical. The reason is that if firms anticipate that their production line might be made obsolete at some point in the future, they will tend to base their decisions to innovate on the current state of profits. If profits fall substantially during recessions, as empirical evidence suggests they do, entrepreneurs might very well undertake more innovation precisely when it is suboptimal to do so. To demonstrate this formally, I had to impose convenient assumptions that make it possible to solve for an equilibrium in the presence of fixed costs analytically. As such, it is hard to use the model to assess whether fixed costs in practice are sufficiently large to account for the fact that innovation is procyclical despite clear incentives to concentrate it in booms. Likewise, it is hard to ascertain the magnitude of the welfare cost of this inefficiency. Incorporating the elements described here in a more realistic model that can be used to address such questions, which would ultimately require an inherently numerical approach, seems like a natural direction for future work.

3. Implications for the Cost of Fluctuations

To illustrate the importance of the discrepancy between the timing of optimal innovation and equilibrium innovation, I now show that the solution to the planner's problem and the decentralized equilibrium can lead to substantially different welfare conclusions concerning the impact of cyclical fluctuations and the desirability of stabilization. In particular, I show that if we focus on the planner's problem in the above economy, we would conclude that cyclical fluctuations always allow the agent to achieve a (weakly) higher utility. Thus, to the extent that fluctuations in productivity arise from factors that can be affected by policy, e.g. sunspots that can be defused by threats to carry out certain policies if agents undertake particular choices, it would be undesirable to use stabilization policy to offset these fluctuations. However, if we consider the equilibrium of the decentralized economy, we would conclude that fluctuations might actually lower welfare by raising the cost of growth relative to the stable economy. In this case, stabilization might be desirable after all. Thus, the discrepancy between equilibrium and optimal innovation can lead to very different conclusions regarding business cycles and macroeconomic policy.

Formally, consider two environments, one in which productivity is constant and equal to $\overline{Z}$, and the other in which productivity follows a Markov process as described above but with the same unconditional average productivity, i.e.

$$E[Z_t] = \tfrac{1}{2}(Z_0 + Z_1) = \overline{Z}$$

For the stochastic environment, suppose initial productivity is distributed with equal probability over Z_0 and Z_1, and that the initial generation of technologies used in producing intermediate goods is the same as in the nonstochastic environment. Both economies thus have the same ability to convert resources into consumption goods over the long-run; the only difference is that the ability to produce consumption goods fluctuates in one environment but not the other. We can measure the welfare impact of fluctuations by comparing the utility of the representative agent in the two environments. If utility is higher in the nonstochastic environment, exogenous fluctuations should be viewed as imposing a social burden on agents. Following Lucas (1987), we could quantify this measure using compensating variation in consumption units, i.e. by determining how much consumption per unit of time we would have to give agents in the volatile environment to leave them as well off as in the stable environment.[10]

Turning first to the social optimum, note that a policy which keeps innovation constant over time will yield the same expected output $E\left[Z_t\left(\lambda^{M_t}(L-R)\right)^{1-\alpha}\right]$ in the two environments, since in both cases M_t follows the same deterministic path, and in both cases $E(Z_t) = \overline{Z}$. Since a benevolent social planner will choose a constant level of innovation when $Z_t = \overline{Z}$ for all t, it follows that the social planner can deliver the same unconditional expected utility for the agent in the stochastic environment as in the nonstochastic environment. But since Proposition 3 implies the planner deliberately varies innovation over the cycle (at least in an interior optimum), it follows that the planner can do strictly better in the presence of fluctuations than when productivity is constant. This captures the intuition implicit in the recent view that business cycles can play a beneficial, even welfare-enhancing, role. The reason is that when productivity in one sector oscillates between high and low levels, agents can specialize their activity to the sector that is relatively more productive. The drawback of this scheme is that in the absence of savings, the agent must tolerate consumption volatility. But as long as agents are not too risk averse, society will be strictly better off with fluctuations than without. Even if agents are risk-averse, allowing for the possibility of storage might potentially allow a planner to achieve growth at a lower cost without forcing a highly volatile consumption stream on the agent, and thus cycles might raise welfare in this case as well.

In the decentralized equilibrium for the same economy, however, it need not be true that fluctuations are welfare improving. This is easiest to see when fixed costs are large enough that equilibrium innovation is procyclical. In that case, resources are systematically allocated to the wrong sector. That is, rather than shifting resources to the sector that is relatively more productive at a given instant, they are shifted to the sector that is relatively less productive. As such, it is possible that introducing fluctuations will lower welfare, even though fluctuations allow the agent to attain a higher utility. In fact, for the parameters in Table 1, one can confirm that the unconditional expected utility of the agent in equilibrium is lower in the environment with fluctuations than in the environment where productivity if constant over time. This will not necessarily be true for all parameter values. In particular, recall that growth in the economy described above can be inefficient in two ways: the average level of innovation might be too high or too low, and the timing of innovation might be suboptimal. While fluctuations give rise

[10] It is important to note that this thought experiment captures the effect of fluctuations in productivity that are exogenous to the innovation process. By contrast, Shleifer (1986), Francois and Shi (1999), Francois and Lloyd-Ellis (2002), and Jovanovic (2002) consider the welfare effects of fluctuations in innovation that arise endogenously. Their welfare calculations are thus not directly comparable to those described here.

to an inefficiency in the timing of innovation that would not arise in the absence of fluctuations, they can also affect the level of innovation in a way that serves to increase welfare. As a result, the economy with fluctuations might be associated with higher welfare even when equilibrium innovation is procyclical, or with lower welfare when equilibrium innovation is countercyclical so that agents do engage in intertemporal substitution.

Note that in the case where fluctuations reduce welfare in equilibrium, business cycles are costly even though the agent is risk-neutral. This distinguishes the cost of fluctuations above from the usual cost that is due to an aversion to consumption volatility, as described in Lucas (1987). Similarly, the cost here is distinct from the cost described in Barlevy (2002), which is also due to an underlying concavity, not in utility but in the function that maps investment into growth. In the latter case, fluctuations are costly because they lead to lower growth from a given average level of investment due to a Jensen's inequality effect. In both cases, then, fluctuations are costly because they introduce variance. By contrast, the inefficient timing of innovation implies fluctuations are costly because they introduce a particular covariance between the amount of resources employed in a sector and that sector's productivity. Variance in itself in not costly in my model, and in fact, if the amount of resources employed in a sector are positive correlated with productivity, volatility can be welfare enhancing.

Interestingly, the above cost of fluctuations can potentially be avoided without having to stabilize the underlying source of fluctuations. Since the cost of fluctuations rests on inefficient timing, one can avoid these costs by taxing and subsidizing innovation activity (at rates which vary with the exogenous component of productivity Z_t) in a way that eliminates the underlying inefficiency.[11] Thus, even if the government has no ability to directly affect the underlying source of fluctuations, there is still a way for it to avoid the particular cost of business cycles described above. This again is different from costs that are due to the volatility of cycles and which cannot be avoided in inherently volatile environments.

4. Extensions

Although the results above are quite intuitive, the mathematical analysis involved in formalizing them is involved. This limits my ability to modify the basic structure of the model to allow for more realistic complications without sacrificing analytical tractability. Nonetheless, I can

[11] Canton and Uhlig (1998) characterize the optimal tax policy in a similar model with no fixed costs of production.

still speculate on directions the model can be modified to address some of its shortcomings and whether they are likely to affect its main results. This section briefly comments on several such extensions.

One problem with the model is that in practice, innovations often take time to develop to fruition. Firms should therefore not expect to be able to bring their innovations on-line as soon as aggregate conditions improve, which is technically possible under the Poisson technology for innovation in the model. Introducing a time-to-build feature could make entrepreneurs reluctant to undertake more innovation in booms if they are inherently temporary. Formally, suppose an idea discovered at date t can only be used for production at date $t+T$. Solving for either the socially optimal path or a symmetric equilibrium in this case is quite complicated, because the relevant state of the world now involves a continuum of productivity levels $[Z_{t-T}, Z_t]$ as opposed to just the current level of productivity Z_t. Even without solving the model, we can use continuity and the results for $T=0$ to establish that inefficient timing would continue to arise for small T. For large T, the same intuition as before suggests that the socially optimal path and the equilibrium path are likely to remain asynchronized when profits are sufficiently volatile. On the one hand, even with delay, there is a clear incentive for the planner to allocate more labor to production in periods of high productivity, although the precise path will now be more complicated given that a part of productivity growth over the next T units of time is already predetermined by decisions made before date t. At the same time, the persistence of the stochastic process for Z_t implies that $\Pr(Z_{t+T} = Z_t \mid Z_t) > \Pr(Z_{t+T} \neq Z_t \mid Z_t)$, so a potential entrant who is constrained to implement his invention only after T units of time still expects profits to be higher at the time of implementation if current profits are high than if they are low. As long as κ is sufficiently large, the expected value of an innovation should still rise disproportionately with current productivity Z_t.

Another feature of the model that might be questionable is its assumption that firms implement innovations as soon as they discover them. In particular, as originally pointed out by Shleifer (1986), entrepreneurs might benefit from strategically delaying implementation of new technologies. This is even more relevant when productivity fluctuates exogenously, since firms can potentially undertake innovation in recessions when the cost of R&D is low but wait until booms to implement them and capture the high profits available at these times. While this intuition is suggestive, it is not necessarily true that allowing for strategic delay will eliminate the inefficient timing of innovation. First and foremost, allowing for strategic delay of innovations typically gives rise to multiple equilibria. It is therefore not obvious that an equilibrium with delay will displace the equilibrium in Proposition 4 rather than represent an additional equilibrium. For certain parameter restrictions, we can in fact confirm that the equilibrium in

Proposition 4 remains an equilibrium even with the possibility of strategic delay:

Proposition 5: Suppose $\lambda < e^{1-\alpha}$ and $\frac{\mu}{\rho+\mu} < \frac{Z_0}{Z_1}$. Then there exists a $\kappa' > \kappa^*$ as defined in Lemma 2 such that the (R_0, R_1) identified in Proposition 4 remains an equilibrium for all $\kappa \in (\kappa^*, \kappa')$ even if agents can delay implementation.

Intuitively, as long as the discount rate ρ is large and regime switches are rare so that μ is small, it will not pay to delay innovation until a boom given the long expected wait until it arrives. If the assumptions of Proposition 5 are not satisfied, firms might prefer to delay implementation until aggregate productivity is high, and it is not obvious that the equilibrium identified in Proposition 4 survives. However, even if equilibrium innovation is countercyclical, the timing of actual innovation is likely to remain inefficient in equilibrium. In particular, even if equilibrium innovation were countercyclical, the inverted timing could simply shift from innovation decisions to implementation decisions. That is, since delay is never socially efficient, the planner would choose paths that concentrate both innovation and implementation in recessions, whereas implementation in equilibrium would be delayed until booms. The inefficient timing of implementation as opposed to innovation is related to results in Caballero and Hammour (1996), who also argue that the timing of the adoption of new technologies can be distorted in the presence of frictions. However, analyzing equilibrium in the case of strategic delay requires a different analysis from the one used in proving previous propositions and is a far more complicated task.[12]

Finally, another weakness of the model concerns its failure to match empirical evidence on job reallocation over the business cycle. In the model, the only instance in which resources are reallocated from one production site to another is when a firm succeeds in innovation and displaces an incumbent monopolist. Thus, if innovation is procyclical in equilibrium as implied by the model, job reallocation should be similarly procyclical. Empirically, Davis and Haltiwanger (1992) report that job reallocation is countercyclical, at least in the manufacturing sector. The model therefore needs to be modified in some way to account for the fact that recessions are

[12] Some of the complications emerge from the recent work of Francois and Lloyd-Ellis (2002), who study delay in a similar model where productivity in the final goods sector is constant. They show that since implementation allows rivals to begin working on the next generation technology, there might be an incentive for entrepreneurs to delay implementation. A key question is whether firms can credibly signal their success to deter rivals from also innovating between when a firm succeeds in innovation and when it actually implements that innovation. Francois and Lloyd-Ellis show that there exists an equilibrium in which signals are credible, but this requires that there are some intermediate goods that have yet to be innovated. Their results also suggest that as long as innovation is bounded away from zero, within a finite amount of time all goods will be innovated. In the case of shocks, firms might wish to delay implementation until productivity is high. But since this can take an arbitrarily long period of time, we cannot assume there is always an intermediate good that has yet to be innovated.

associated with greater job reallocation. This can potentially be addressed by introducing an additional source of reallocation that is not related to innovation activity. For example, following Young (1998) or Aghion and Saint Paul (1998), we can introduce a "horizontal" variety dimension in addition to the "vertical" quality ladder dimension that is already present in the model. That is, suppose labor is used to produce capital goods, where the number of different capital goods N is endogenous, and these capital goods can together be converted into a unit mass of intermediate goods according to the same linear specification as before, where each intermediate good j can be produced more efficiently under successively higher generations. If the number of capital goods N decreases with productivity, say because fewer capital goods can be profitably produced at lower productivity, recessions could be associated with increased job reallocation despite the decline in innovation activity. Specifically, a recession would trigger a burst of job destruction as the variety of capital goods declines. Eventually, though, measured job reallocation would decline as innovative activity is diminished and fewer resources are shifted between older and newer production technologies. By contrast, as long as we introduced some friction in the process for creating new capital goods, such as search costs or convex creation costs as in Caballero and Hammour (1996), a boom would not lead to a similar burst of job creation, and reallocation of labor among production sites for different capital goods would appear countercyclical. This pattern is consistent with recent evidence in Caballero and Hammour (1999). They argue that while recessions are associated with an initial increase in job reallocation, they eventually lead to decline in cumulative job reallocation, as would be implied by the decline in innovation activity.

5. Conclusion

The purpose of this paper is to explore the behavior of innovation in response to cyclical fluctuations that affect the relative productivity in the goods sector and the innovation sector. Using a Schumpeterian growth model based on the model of Grossman and Helpman (1991), I demonstrate that fluctuations can give rise to an inefficient timing of innovation activity in which agents shift less innovation to recessions than would be optimal. Quite strikingly, it is even possible that agents will concentrate innovation in booms. This can explain why R&D expenditures appear to be procyclical, even though productivity considerations clearly suggest they should by countercyclical.

The reason for this distorted timing is the creative destruction aspect inherent to Schumpeterian growth models. That is, since firms anticipate that their production technology will eventually be rendered obsolete, they discount profits that accrue in the more distant future too heavily and instead base their decisions on current economic conditions. If this were indeed the reason why equilibrium R&D is procyclical, we would anticipate that activities that similarly come at the expense of production but which allow agents to capture all of the returns to their investments should be concentrated in recessions. An obvious example of such an activity is human capital accumulation. Like R&D, education and training require sacrificing labor services. In contrast to R&D, though, an agent that accumulates human capital does not run a risk of having his investment stolen away by rival workers. Thus, there is potentially a similar argument for concentrating human capital accumulation in recessions. That said, empirical evidence suggests the productivity of skill accumulation may not be constant over the cycle. For example, DeJong and Ingram (2001) offer evidence that the productivity of skill accumulation is positively correlated with productivity in the goods sector. But they also find that this correlation is not enough to overturn the prediction that investment in skills should be countercyclical, and they offer evidence that investment in human capital is countercyclical. Likewise, data on college enrollments reveals a markedly countercyclical pattern in schooling, as documented in Betts and McFarland (1995) and Dellas and Sakellaris (1997). This suggests that the creative destruction aspect of R&D is more likely to account for the procyclical pattern in R&D than a conceptual flaw in the intertemporal substitution hypothesis.

The results here are related in spirit to other recent work that has also questioned the notion that recessions contribute positively to enhancing future economic activity. In the particular context of R&D examined here, intertemporal substitution considerations suggest that it will be optimal to take advantage of downturns to promote growth-enhancing activities, and that because of this fluctuations can play a welfare-enhancing role. In practice, though, there are frictions that arise in market environments that blunt these incentives. Thus, although recessions might encourage certain types of productivity-enhancing activities that may benefit the economy in the long run, they also might suppress other activities in a way that lowers overall welfare. While the analysis here helps to explain *why* R&D can be procyclical even when technological considerations would dictate otherwise, further work is necessary to gauge the empirical plausibility of the explanation it offers, and to what extent this failure contributes to the costs of business cycles and necessitates some form of policy intervention.

Appendix

Proof of Propositions 1 and 3: For given values of $\{R_i\}_{i=0,1}$, the system given by (1.11) reduces to ordinary linear differential equations in $V(Z_i, M)$. Standard theorems ensure this system has a unique solution. Hence, starting with values for R_i, we can use the method of undetermined coefficients to find the unique value functions $V(Z_i, M)$ associated with a given pair (R_0, R_1). I conjecture that the value function $V(\cdot,\cdot)$ takes the form

$$V(Z_i, M) = v_i \lambda^{M(1-\alpha)}$$

Differentiating this function with respect to M yields

$$\frac{\partial V}{\partial M} = (1-\alpha) v_i \lambda^{M(1-\alpha)} \ln \lambda$$

which simplifies the differential equations above to a system of independent linear equations in the coefficients v_i:

$$\rho v_i = Z_i (L - R_i)^{1-\alpha} - \kappa + \mu (v_{-i} - v_i) + (1-\alpha) v_i \phi R_i \ln \lambda$$

This yields a unique solution (v_0, v_1) as functions of (R_0, R_1).

Since the RHS of (1.11) is strictly concave in R_i, the first order condition is both necessary and sufficient to characterize the optimal R_i. The first order condition is given by

$$-(1-\alpha) Z_i \lambda^{M(1-\alpha)} (L - R_i)^{-\alpha} + \frac{\partial V}{\partial M}\phi \leq 0 \tag{5.1}$$

with equality if $R_i > 0$. Substituting the expression for $V(\cdot,\cdot)$, we obtain

$$R_i = \begin{cases} L - \left(\dfrac{Z_i}{v_i \phi \ln \lambda}\right)^{\frac{1}{\alpha}} & \text{if } v_i > \dfrac{Z_i}{\phi L^{\alpha} \ln \lambda} \\ 0 & \text{else} \end{cases} \tag{5.2}$$

If we substitute this expression into the asset equation (1.11), we obtain a pair of equations with v_{-i} as a function of v_i that hold at the optimal R_i:

$$v_{-i} = f_{-i}(v_i) = \begin{cases} \dfrac{(\rho + \mu - (1-\alpha)\phi L \ln \lambda)}{\mu} v_i - \dfrac{\alpha}{\mu} Z_i^{\frac{1}{\alpha}} (v_i \phi \ln \lambda)^{1-\frac{1}{\alpha}} + \dfrac{\kappa}{\mu} & \text{if } v_i > \dfrac{Z_i}{\phi L^{\alpha} \ln \lambda} \\ \dfrac{\rho + \mu}{\mu} v_i - \dfrac{Z_i L^{1-\alpha}}{\mu} + \dfrac{\kappa}{\mu} & \text{else} \end{cases}$$

The optimal program corresponds to any pair (v_0^*, v_1^*) which solves the equations

$$\begin{aligned} v_1^* &= f_1(v_0^*) \\ v_0^* &= f_0(v_1^*) \end{aligned}$$

The function $f_{-i}(\cdot)$ is continuous and differentiable, since the left and right hand derivatives at $v_i = \dfrac{Z_i}{\phi L^{\alpha} \ln \lambda}$ are both equal to $\dfrac{\rho + \mu}{\mu}$. Since $\rho > (1-\alpha)\phi L \ln \lambda$, it follows that $\dfrac{\partial f_{-i}(v_i)}{\partial v_i} > 1$ for all v_i. The functions $f_{-i}(\cdot)$ are illustrated in Figure A1, suggesting that there is a unique solution (v_0^*, v_1^*). To establish this formally, I use the fact $\dfrac{df_{-i}}{dv_i} > 1 > 0$ for all v_i implies $f_{-i}(\cdot)$ is invertible. An equilibrium therefore involves a value v_0^* such that $f_1(v_0^*) - f_0^{-1}(v_0^*) = 0$. Differentiating this condition with respect to v_0^* yields

$$\frac{d}{dx}\left[f_1(x) - f_0^{-1}(x)\right] = \frac{df_1}{dx} - \left(\frac{df_0}{dx}\right)^{-1} > 0$$

This monotonicity insures there is at most one value of v_0^*. To establish existence, note that $f_1(0) < 0$ while $f_0^{-1}(0) > 0$. Hence, $f_1(0) - f_0^{-1}(0) < 0$, and is finite. The fact that $\lim_{x\to\infty} \dfrac{df_1}{dx} > 1 > \lim_{x\to\infty}\left(\dfrac{df_0}{dx}\right)^{-1}$ implies $\dfrac{\partial}{\partial x}\left[f_1(x) - f_0^{-1}(x)\right]$ is strictly bounded away from 0, and so $f_1(x) - f_0^{-1}(x) \to \infty$ as $x \to \infty$. The existence of v_0^* follows from continuity. This implies there is a unique social solution to the social planner's problem.

Next, suppose that the optimal path dictates $R_i > 0$ for both i. I need to show $R_0 > R_1$. The proof proceeds in two steps. First, I argue that $v_1^* > v_0^*$. Since $R_i > 0$, the asset equations imply

$$\begin{aligned} v_{-i}^* &= \frac{(\rho + \mu - (1-\alpha)\phi L \ln\lambda)}{\mu} v_i^* - \frac{\alpha}{\mu} Z_i^{\frac{1}{\alpha}} \left(v_i^* \phi \ln\lambda\right)^{1-\frac{1}{\alpha}} + \frac{\kappa}{\mu} \\ &\equiv a v_i^* - b Z_i^{\frac{1}{\alpha}} \left(v_i^*\right)^{1-\frac{1}{\alpha}} + \frac{\kappa}{\mu} \end{aligned}$$

Consider the fixed point $\widehat{v}_i$ which solves

$$\widehat{v}_i = a\widehat{v}_i - b Z_i^{\frac{1}{\alpha}} \left(\widehat{v}_i\right)^{1-\frac{1}{\alpha}} + \frac{\kappa}{\mu}$$

It is easy to show $\widehat{v}_i$ exists and is unique. Implicit differentiation implies

$$\frac{d\widehat{v}_i}{dZ_i} = \frac{\frac{b}{\alpha}\left(\frac{\widehat{v}_i}{Z_i}\right)^{1-\frac{1}{\alpha}}}{(\alpha - 1) + \frac{1-\alpha}{\alpha} b \left(\frac{Z_i}{\widehat{v}_i}\right)^{\frac{1}{\alpha}}} > 0$$

so that $Z_0 < Z_1 \Rightarrow \widehat{v}_0 < \widehat{v}_1$. Since $\dfrac{df_i^{-1}}{dx} < 1$, we know that for any $x < \widehat{v}_1$, it follows that $x - f_0^{-1}(x) < 0$. Hence,

$$\begin{aligned} f_1(\widehat{v}_0) - f_0^{-1}(\widehat{v}_0) &= \widehat{v}_0 - f_0^{-1}(\widehat{v}_0) \\ &< 0 \end{aligned}$$

where the inequality uses the fact that $\widehat{v}_0 > \widehat{v}_1$. Since $f_1(v_0^*) - f_0^{-1}(v_0^*) = 0$ and $f_1(x) - f_0^{-1}(x)$ is increasing in x, it follows that $v_0^* > \widehat{v}_0$. But since $\dfrac{df_i^{-1}}{dx} > 1$, the fact that $f_1(\widehat{v}_0) = \widehat{v}_0$ implies $f_1(x) > x$ for any $x > \widehat{v}_0$. Hence, $f_1(v_0^*) > v_0^*$. But since $v_1^* = f_1(v_0^*)$, it follows that $v_1^* > v_0^*$.

Next, I use the fact that $v_1^* > v_0^*$ to argue $\dfrac{v_1}{Z_1} < \dfrac{v_0}{Z_0}$, which is sufficient to establish $R_1 < R_0$ from the first-order condition above. Combining the equations $v_{-i}^* = f_{-i}(v_i^*)$ for both values yields the equation

$$a v_0^* - b Z_0^{\frac{1}{\alpha}} \left(v_0^*\right)^{1-\frac{1}{\alpha}} - v_1^* = a v_1^* - b Z_1^{\frac{1}{\alpha}} \left(v_1^*\right)^{1-\frac{1}{\alpha}} - v_0^*$$

which can be rearranged to yield

$$\frac{v_0^*}{v_1^*} = \frac{(a+1) - b\left(\frac{Z_1}{v_1^*}\right)^{\frac{1}{\alpha}}}{(a+1) - b\left(\frac{Z_0}{v_0^*}\right)^{\frac{1}{\alpha}}}$$

so that

$$v_1^* > v_0^* \Leftrightarrow \frac{v_1^*}{v_0^*} < \frac{Z_1}{Z_0}$$

But given the expression for R_i in (5.2), this implies $R_0 > R_1$.

Finally, suppose $R_i = 0$ for some i. The proposition follows if we can rule out the case where $R_0 = 0$ and $R_1 > 0$ for sufficiently small κ. Once again, it will be enough to show that $\frac{v_1^*}{v_0^*} < \frac{Z_1}{Z_0}$. Let $v_i^*(Z_0, Z_1)$ denote the values of v_i^* given Z_0 and Z_1. It will be enough to prove that

$$\frac{\partial}{\partial Z_1}\left[\frac{Z_0 v_1^*(Z_0, Z_1)}{Z_1 v_0^*(Z_0, Z_1)}\right] < 0 \tag{5.3}$$

This is because by integrating (5.3) with respect to Z_1, we obtain

$$\begin{aligned}\frac{Z_0 v_1^*(Z_0, Z_1)}{Z_1 v_0^*(Z_0, Z_1)} &= \left[\frac{Z_0 v_1^*(Z_0, Z_0)}{Z_0 v_0^*(Z_0, Z_0)}\right] + \int_{Z_0}^{Z_1} \frac{\partial}{\partial Z_1}\left[\frac{Z_0 v_1^*(Z_0, Z_1)}{Z_1 v_0^*(Z_0, Z_1)}\right] dZ_1 \\ &< \left[\frac{Z_0 v_1^*(Z_0, Z_0)}{Z_0 v_0^*(Z_0, Z_0)}\right] = 1\end{aligned}$$

Note that (5.3) holds if and only if

$$\frac{\partial}{\partial Z_1} \ln\left(\frac{Z_0 v_1^*(Z_0, Z_1)}{Z_1 v_0^*(Z_0, Z_1)}\right) < 0$$

or alternatively if

$$\frac{\partial v_1^*/\partial Z_1}{v_1^*/Z_1} < 1 + \frac{\partial v_0^*/\partial Z_1}{v_0^*/Z_1}$$

Differentiating the asset equations with respect to Z_1 yields

$$\begin{aligned}\frac{\partial v_1}{\partial Z_1} &= \begin{cases} \dfrac{(L-R_1)\phi \ln\lambda}{\phi(L-R_1)\ln\lambda + \alpha\phi R_1 \ln\lambda} \dfrac{v_1}{Z_1} & \text{if } R_1 > 0 \\[2ex] \dfrac{(\rho+\mu) Z_0 L^{1-\alpha}}{(\rho+\mu) Z_1 L^{1-\alpha} + \mu(Z_0 L^{1-\alpha} + A) - (\rho+2\mu)\kappa} \dfrac{v_1}{Z_1} & \text{if } R_1 = 0 \end{cases} \\ \frac{\partial v_0}{\partial Z_1} &= \frac{\mu}{\rho+\mu-(1-\alpha)\phi R_{-i}\ln\lambda} \frac{\partial v_1}{\partial Z_1}\end{aligned}$$

where A is defined by

$$Z_0 L^{1-\alpha} + A = \max_{R_0}\left\{Z_0 (L-R_0)^{1-\alpha} + (1-\alpha) v_0 \phi R_0 \ln\lambda\right\}$$

so that $A \geq 0$. Provided $\kappa < \frac{\mu}{\rho+2\mu} Z_0 L^{1-\alpha}$, $\frac{\partial v_1}{\partial Z_1} < \frac{v_1}{Z_1}$, and $\frac{\partial v_0}{\partial Z_1} > 0$. This insures $\frac{\partial v_1^*/\partial Z_1}{v_1/Z_1} < 1 < 1 + \frac{\partial v_0^*/\partial Z_1}{v_0^*/Z_1}$, completing the proof. ■

Lemma 1: Let $h(\xi) = \frac{\mu}{\omega(L)}\left[\frac{Z_1}{Z_0}\xi^{\alpha} - \frac{Z_0}{Z_1}\xi^{1-\alpha}\right]$, where $\omega(L) = \rho + \mu + (1-(1-\alpha)\ln\lambda)\phi L$. There exists a unique $\xi^* > 0$ such that $1 - \xi^* \gtreqless h(\xi^*)$ if $\xi \lesseqgtr \xi^*$. This unique solution ξ^* lies in the interval $(0,1)$.

Proof of Lemma 1: First, I claim there exists a $\xi^* \in (0,1)$ for which $1 - \xi = h(\xi)$. This is straightforward: if $\xi = 0$, we have

$$1 - \xi = 1 > 0 = h(\xi)$$

while if $\xi = 1$, we have

$$1 - \xi = 0 < \frac{\mu}{\omega(L)}\left[\frac{Z_1}{Z_0} - \frac{Z_0}{Z_1}\right] = h(\xi)$$

where the inequality relies on the fact that $\omega(L) > 0$ given that $\rho > (1-\alpha)\phi L \ln\lambda$. The claim follows from continuity.

To prove ξ^* is unique, I proceed in two steps. Differentiating $h(\cdot)$ yields

$$h'(\xi) = \frac{\mu}{\omega(L)}\left[\alpha\frac{Z_1}{Z_0}\xi^{\alpha-1} - (1-\alpha)\frac{Z_0}{Z_1}\xi^{-\alpha}\right]$$

For $\xi \geq 1$, we have

$$\begin{aligned} \alpha\frac{Z_1}{Z_0}\xi^{\alpha-1} - (1-\alpha)\frac{Z_0}{Z_1}\xi^{-\alpha} &> -(1-\alpha)\frac{Z_0}{Z_1}\xi^{-\alpha} \\ &> -1 \end{aligned}$$

Since at $\xi = 1$, $h(\xi) > 1 - \xi$, a necessary condition for there to exist a $\xi^* > 1$ such that $1 - \xi^* = h(\xi^*)$ is that there exists a $\xi > 1$ such that $h'(\xi) < -1$. Thus, there exists no $\xi > 1$ for which $1 - \xi = h(\xi)$.

Next, I need to show there is a unique $\xi^* \in (0,1)$ for which $1 - \xi^* = h(\xi^*)$. Consider first the case where $\alpha > \frac{1}{2}$. Differentiating $h(\cdot)$ establishes that $h'(\xi) \geq 0$ if and only if

$$\frac{\alpha}{1-\alpha}\left(\frac{Z_1}{Z_0}\right)^2 > \xi^{1-2\alpha}$$

For $\alpha > \frac{1}{2}$, $1 - 2\alpha < 0$, and so $h'(\xi)$ is negative if $0 < \xi < \left[\frac{\alpha}{1-\alpha}\left(\frac{Z_1}{Z_0}\right)^2\right]^{\frac{1}{1-2\alpha}}$ and positive if $\xi > \left[\frac{\alpha}{1-\alpha}\left(\frac{Z_1}{Z_0}\right)^2\right]^{\frac{1}{1-2\alpha}}$. Since $h(0) = 0$, it follows that $h(\xi) < 0$ for $0 < \xi < \left[\frac{\alpha}{1-\alpha}\left(\frac{Z_1}{Z_0}\right)^2\right]^{\frac{1}{1-2\alpha}}$.

Hence, $1-\xi > 0 > h(\xi)$ for $\xi < \left[\frac{\alpha}{1-\alpha}\left(\frac{Z_1}{Z_0}\right)^2\right]^{\frac{1}{1-2\alpha}} < 1$. Since $h'(\xi)$ is strictly positive for $\left[\frac{\alpha}{1-\alpha}\left(\frac{Z_1}{Z_0}\right)^2\right]^{\frac{1}{1-2\alpha}}$, it follows that ξ^* is unique and $1-\xi^* \gtreqless h(\xi^*)$ if $\xi \lesseqgtr \xi^*$

If $\alpha = \frac{1}{2}$, $h(\xi)$ simplifies to $\frac{\mu}{\omega(L)}\left[\frac{Z_1 - Z_0}{Z_1 Z_0}\right]\xi^{\frac{1}{2}}$ which is monotonically increasing in ξ while $1-\xi$ is monotonically decreasing. This again insures ξ^* is unique and $1-\xi^* \gtreqless h(\xi^*)$ if $\xi \lesseqgtr \xi^*$

Finally, if $\alpha < \frac{1}{2}$, it is enough to prove that $h'(\xi) > -1$ for all $\xi \in (0,1]$. Differentiating $h(\cdot)$ twice yields

$$h''(\xi) = \frac{\alpha(1-\alpha)\mu}{\omega(L)}\left[\frac{Z_0}{Z_1}\xi^{-\alpha-1} - \frac{Z_1}{Z_0}\xi^{\alpha-2}\right]$$

so that

$$h''(\xi) \gtreqless 0 \Leftrightarrow \xi \gtreqless \left(\frac{Z_1}{Z_0}\right)^{\frac{2}{1-2\alpha}}$$

Thus, the derivative attains a minimum at $\xi = \left(\frac{Z_1}{Z_0}\right)^{\frac{2}{1-2\alpha}}$ which for $\alpha < \frac{1}{2}$ is strictly greater than 1. But it was previously argued that $h'(\xi) > -1$ for all $\xi \geq 1$. Hence, $h'(\xi) > -1$ for all $\xi \in (0,\infty)$ if $\alpha < \frac{1}{2}$.

Lastly, since at $\xi = 0$, $1-\xi = 1 > 0 = h(\xi)$, continuity implies $1-\xi > h(\xi)$ for all $\xi < \xi^*$. Likewise, at $\xi = 1$, $1-\xi = 0 < h(1)$, so by continuity it follows that $1-\xi < h(\xi)$ for $\xi > \xi^*$. This establishes the lemma. ■

Proof of Proposition 2: Since $R_i \geq 0$, the case where $R_1 = 0$ trivially satisfies the claim. If $R_0 = 0$, we need to verify that $R_1 = 0$. To show this, suppose not, i.e. suppose $R_1 > 0 = R_0$. Then it follows that $\phi v_1 = 1 \geq \phi v_0$. Substituting in for v_i, we get

$$\left\{\begin{array}{l}\omega(R_0)(L-R_1) + \\ \quad \mu\frac{Z_0}{Z_1}(L-R_0)^{1-\alpha}(L-R_1)^{\alpha}\end{array}\right\} \geq \left\{\begin{array}{l}\omega(R_1)(L-R_0) + \\ \quad \mu\frac{Z_1}{Z_0}(L-R_1)^{1-\alpha}(L-R_0)^{\alpha}\end{array}\right\}$$

Since $v_1(R_1, R_0) = 0$ if $R_1 = L$, then $R_1 < L$ in any such equilibrium. This allows us to define ξ such that

$$R_0 = \xi R_1 + (1-\xi)L$$

Note that since $R_1 < L$, by construction, $\xi \geq 0$, and $R_0 > R_1$ implies $\xi \in [0,1)$ while $R_0 < R_1$ implies $\xi > 1$. After substituting in for $\omega(R)$ and rewriting R_0 in terms of R_1, we can rewrite the inequality $v_1 \geq v_0$ in terms of ξ:

$$1-\xi \geq \frac{\mu}{\omega(L)}\left[\frac{Z_1}{Z_0}\xi^{\alpha} - \frac{Z_0}{Z_1}\xi^{1-\alpha}\right]$$

Applying lemma 1, it follows that $\xi < \xi^* < 1$, which implies $R_0 > R_1$, a contradiction. Thus, $R_0 = 0$ implies $R_1 = 0$.

Finally, if R_1 and R_0 are both positive, it must be true that $v_1 = v_0 = \frac{1}{\phi}$, which in turn implies that $1 - \xi = h(\xi)$. But from the lemma, the unique ξ^* which satisfies this equation is less than 1, which implies $R_0 \geq R_1$.

Next, I show that there exists a unique symmetric Markov-perfect equilibrium when $\lambda < e^{\frac{1}{1-\alpha}}$. This condition implies that $(1-\alpha)\ln\lambda < 1$, which implies

$$\omega'(R) = (1 - (1-\alpha)\ln\lambda)\,\phi > 0$$

Using (1.20), the fact that $\omega'(R) > 0$ can be shown to imply that $\frac{\partial v_i}{\partial R_i} < 0$ and $\frac{\partial v_i}{\partial R_{-i}} < 0$. Recall from the proof of Proposition 2 that in any Markov-perfect equilibrium in which $\phi v_i(R_i, R_{-i}) = 1$ for both i, the levels of innovation for the two levels of productivity are related by $R_0 = \xi^* R_1 + (1-\xi^*)\,L$, where ξ^* is a constant. Hence, there can be at most one equilibrium in which $\phi v_0 = \phi v_1 = 1$. For suppose there were two such equilibria, $(R_0, R_1) \neq (R_0', R_1')$ where wlog $R_0' > R_0$. Since ξ^* is constant, it follows that $R_1' > R_1$, but since v_i is decreasing in both R_i and R_{-i} it is impossible that $v_i(R_i, R_{-i}) = v_i\left(R_i', R_{-i}'\right) = \frac{1}{\phi}$.

If $(R_0, R_1) = (0, 0)$ is an equilibrium, given that $\frac{\partial v_i}{\partial R_i} < 0$, it follows that for any $(R_0, R_1) \neq 0$ there always exists some $i \in \{0, 1\}$ such that $\phi v_i < 1$ but $R_i > 0$, which is inconsistent with equilibrium. In this case, $(0, 0)$ would be the unique equilibrium. Without loss of generality, then, I henceforth assume that if an equilibrium exists, it is not equal to $(0, 0)$.

I begin by arguing that if there exists an equilibrium (R_0^*, R_1^*) where $\phi v_i\left(R_i^*, R_{-i}^*\right) = 1$ for both i, there exists no other equilibria in which $R_i = 0$ and $\phi v_i(0, R_{-i}) < 1$ for some $i \in \{0, 1\}$. For each i, define the contour sets

$$\Omega_i = \{(R_i, R_{-i}) \mid \phi v_i(R_i, R_{-i}) = 1\}$$

for all values of $(R_0, R_1) \geq (0, 0)$. These sets are illustrated in Figure A2. Using the implicit function theorem and the fact that $\frac{\partial v_i}{\partial R_i}$ and $\frac{\partial v_i}{\partial R_{-i}}$ are both strictly negative, we can establish that the graphs of Ω_i form connected, downward sloping curves in (R_0, R_1) space. If there exists an equilibrium $(R_0^*, R_1^*) \neq (0, 0)$ such that $\phi v_i\left(R_i^*, R_{-i}^*\right) = 1$ for both i, the sets Ω_i must both be nonempty. Since $\frac{\partial v_i}{\partial R_{-i}} < 0$ and $\phi v_i\left(R_i^*, R_{-i}^*\right) = 1$, then $\phi v_i(R_i^*, 0) > 1$. Since $\phi v_i(L, 0) = 0$, there exists an $R_i' \geq 0$ such that $(R_i', 0) \in \Omega_i$ by continuity. Hence, The graph of Ω_i intersects the R_i axis.

Next, define $R_{-i}'' > 0$ as the value of R_{-i} such that $\phi v_i\left(0, R_{-i}''\right) = 1$. If no such value exists, I adopt the convention that $R_{-i}'' = \infty$. I now argue that $R_1'' > R_1'$. The statement follows trivially if $R_1'' = \infty$. If $R_1'' < \infty$, I argue that $\phi v_1(R_1'', 0) > 1$. For suppose not, i.e. suppose $\phi v_1(R_1'', 0) \leq 1$. Since $\phi v_0(0, R_1'') = 1$, it follows that either $(0, R_1'')$ constitutes an equilibrium, or there exists some $R_1''' \in (R_1'', L)$ such that $\phi v_1(R_1''', 0) = 1$, from which it follows that $(0, R_1''')$ is an equilibrium. Since $R_0 \geq R_1$ in any Markov-perfect equilibrium, it follows that $R_1'' = R_1''' = 0$. Since $\phi v_0(R_0^*, R_1^*) = 1$ and $(R_0^*, R_1^*) \neq (0, 0)$ by assumption, $\frac{\partial v_i}{\partial R_i} < 0$ implies $\phi v_0(0,0) > 1$, a contradiction. Hence, $\phi v_1(R_1'', 0) > 1 = \phi v_1(R_1', 0)$. Since v_i is decreasing in R_i, it follows that $R_1'' > R_1'$ as claimed.

The fact that $R_1'' > R_1'$ can be used to establish that $R_0'' > R_0'$ as well. First, though, I argue that at the equilibrium (R_0^*, R_1^*),

$$\left.\frac{dR_1}{dR_0}\right|_{\phi v_1=1} > \left.\frac{dR_1}{dR_0}\right|_{\phi v_0=1}$$

To see this, consider a neighborhood around (R_0^*, R_1^*). Recall that for any admissable (R_0, R_1) where R_0 is defined as $\xi R_1 + (1-\xi)L$ for some $\xi \geq 0$, the proof of Proposition 2 above implies that $v_1(R_1, R_0) > v_0(R_0, R_1)$ if and only if $1-\xi > h(\xi)$, which from Lemma 1 holds if and only if $\xi < \xi^*$. Since for any $\varepsilon > 0$, $R_0^* + \varepsilon = \xi R_1^* + (1-\xi)L$ for some $\xi < \xi^*$, it follows that

$$v_1(R_1^*, R_0^* + \varepsilon) > v_0(R_0^* + \varepsilon, R_1^*)$$

Subtracting $v_0(R_0^*, R_1^*) = v_1(R_1^*, R_0^*) = \frac{1}{\phi}$ from both sides, dividing by ε, and taking the limit as $\varepsilon \to 0$ implies

$$\frac{\partial v_1(R_0^*, R_1^*)}{\partial R_0} \geq \frac{\partial v_0(R_1^*, R_0^*)}{\partial R_0}$$

We can further establish this inequality is strict. This is because if $\frac{\partial v_0(R_1^*, R_0^*)}{\partial R_0} = \frac{\partial v_1(R_0^*, R_1^*)}{\partial R_0}$, the derivative of $1-\xi-h(\xi)$ with respect to ξ would be equal to 0 at $\xi = \xi^*$, which is contradicted by the proof of Lemma 1 above. Similarly, for any $\varepsilon > 0$, Lemma 1 implies

$$v_0(R_0^*, R_1^* + \varepsilon) > v_1(R_1^* + \varepsilon, R_0^*)$$

and by an analogous argument,

$$\frac{\partial v_0(R_0^*, R_1^*)}{\partial R_1} \geq \frac{\partial v_1(R_1^*, R_0^*)}{\partial R_1}$$

Taking into account the fact that $\partial v_i / \partial R_i$ and $\partial v_i / \partial R_{-i}$ are both negative, it follows that

$$\frac{\partial v_0/\partial R_0}{\partial v_0/\partial R_1} > \frac{\partial v_1/\partial R_0}{\partial v_1/\partial R_1}$$

which implies

$$\left.\frac{dR_1}{dR_0}\right|_{\phi v_1=1} = -\frac{\partial v_1/\partial R_0}{\partial v_1/\partial R_1} > -\frac{\partial v_0/\partial R_0}{\partial v_0/\partial R_1} = \left.\frac{dR_1}{dR_0}\right|_{\phi v_0=1}$$

Since there can be only one point at which $\phi v_i(R_i, R_{-i}) = 1$, the two contours sets Ω_i intersect only at (R_0^*, R_1^*), and by continuity it follows that $R_0'' > R_0'$.

With these observations, I can finally establish that there exists no other equilibrium $\left(\widehat{R}_i, \widehat{R}_{-i}\right)$ in which $\widehat{R}_i = 0$ for some $i \in \{0,1\}$ and $\phi v_i\left(0, \widehat{R}_{-i}\right) < 1$. This is because if such an equilibrium existed, by definition it must be true that $\phi v_i\left(0, \widehat{R}_{-i}\right) \leq 1$. Since $\phi v_i\left(0, R''_{-i}\right) = 1$ by definition, monotonicity implies $\widehat{R}_{-i} \geq R''_{-i} > R'_{-i}$. But since $R_{-i} > R'_{-i}$, it follows that $\phi v_{-i}\left(R_{-i}, 0\right) < 1$. For this to be an equilibrium, $\widehat{R}_{-i} = 0$. But since there exists an $\left(R_0^*, R_1^*\right) \neq (0,0)$ such that $\phi v_i\left(R_i^*, R_{-i}^*\right) = 1$, it follows that $\phi v_i\left(0,0\right) > 1$, a contradiction.

Finally, suppose there exists no pair $\left(R_0^*, R_1^*\right)$ such that $\phi v_0\left(R_0^*, R_1^*\right) = \phi v_1\left(R_0^*, R_1^*\right) = 1$. We need to establish that there still exists a unique Markov-perfect equilibrium. Suppose first that $\phi v_0\left(0,0\right) \leq 1$. Then for any $\left(R_0, R_1\right) \geq (0,0)$ where $\left(R_0, R_1\right) \neq (0,0)$, it must be the case that $\phi v_0\left(R_0, R_1\right) < 1$. This implies $R_0 = 0$ in any equilibrium, and since $R_0 \geq R_1$ in any equilibrium according to Proposition 2, $R_0 = R_1 = 0$ must be the unique equilibrium. If we rewrite R_0 as $\xi R_1 + (1-\xi) L$, then $\xi = 1 > \xi^*$. But recall from Lemma 1 that this implies $v_1\left(0,0\right) < v_0\left(0,0\right)$. Since $\phi v_0\left(0,0\right) \leq 1$, it follows that $(0,0)$ is in fact an equilibrium.

This leaves the case where (i) there exists no pair $\left(R_0^*, R_1^*\right)$ such that $\phi v_0\left(R_0^*, R_1^*\right) = \phi v_1\left(R_0^*, R_1^*\right) = 1$ and (ii) $\phi v_0\left(0,0\right) > 1$. By continuity, there exists an $R'_0 < L$ such that $\phi v_0\left(R'_0, 0\right) = 1$. Then I claim $\left(R_0, R_1\right) = \left(R'_0, 0\right)$ is the unique Markov perfect equilibrium. Consider again two cases. First, suppose that $\phi v_1\left(0,0\right) \leq 1$. In that case, monotonicity implies $\phi v_1\left(0, R'_0\right) < 1$ given that $R'_0 > 0$, so that $\phi v_1\left(0, R'_0\right) \leq 1$ and $\left(R'_0, 0\right)$ is indeed an equilibrium. Moreover, since $\phi v_1\left(0,0\right) \leq 1$, then $R_1 = 0$ in any equilibrium, and it follows that $\left(R'_0, 0\right)$ is the unique equilibrium. Lastly, suppose $\phi v_1\left(0,0\right) > 1$. Once again, define R''_0 such that $\phi v_0\left(0, R''_0\right) = 1$, with the convention of setting $R''_0 = \infty$ if no such value exists. We need to show that $R'_0 > R''_0$, which insures $\phi v_1\left(0, R'_0\right) \leq 1$. Suppose not, i.e. suppose $R''_0 \geq R'_0$. But using the same argument as before, we know that $R''_1 > R'_1$. If $R''_0 \geq R'_0$, then by continuity Ω_0 and Ω_1 must intersect, which contradicts the supposition that there exists no solution $\left(R_0^*, R_1^*\right)$. Hence, $\phi v_1\left(0, R'_0\right) \leq 1$, so that $\left(R'_0, 0\right)$ is an equilibrium, and since $R_1 = 0$ in any equilibrium, which insures the equilibrium above is unique.■

Proof of Lemma 2 (in text): Consider the equation $v_0\left(R^*, R^*\right) = v_1\left(R^*, R^*\right)$. It implies

$$\frac{\omega\left(R^*\right) - \mu \dfrac{Z_1}{Z_0}}{(1-\alpha) Z_1 \left(L - R^*\right)^{1-\alpha} - \kappa} - \frac{\omega\left(R^*\right) - \mu \dfrac{Z_0}{Z_1}}{(1-\alpha) Z_0 \left(L - R^*\right)^{1-\alpha} - \kappa} = \frac{\mu}{\lambda}\left[\frac{Z_0}{Z_1} - \frac{Z_1}{Z_0}\right]$$

Differentiate left hand side yields

To establish the existence and uniqueness of R^*, define $y_i = L - R_i$. Using the assumption that $R_i < L$ implies $y_i \neq 0$, we can rearrange the condition $v_0\left(R, R\right) = v_1\left(R, R\right)$ by dividing both sides by y^α and expanding out $\omega\left(L - y\right)$ to obtain

$$\mu\left(Z_1 + Z_0\right)\frac{y^{1-\alpha}}{\kappa} + \frac{\lambda^2}{(\lambda - 1)(1-\alpha)}\left(1 - (1-\alpha)\ln\lambda\right)\phi y = \frac{\lambda^2\left(\omega\left(L\right) + \mu\right)}{(\lambda - 1)(1-\alpha)} \tag{5.4}$$

The LHS of this equation is monotonically increasing in y, and ranges from 0 to ∞ as y ranges from 0 to ∞. Since the RHS above is strictly positive, there exists a unique value y^* for which the equation is satisfied. This translates into a unique value $R^* = L - y^*$.

Note that y^* is monotonically increasing in κ. Taking limits, $y^* \to 0$ as $\kappa \to 0$, while $y^* \to L + \frac{\rho + 2\mu}{(1-(1-\alpha)\ln\lambda)\phi}$ as $\kappa \to \infty$, at which point $\omega(L - y^*) = -\mu$. If we evaluate $v_i(R^*, R^*)$ and substitute in from (5.4), we obtain

$$\begin{aligned} v_i &= (\lambda - 1)\frac{\left(\omega(L-y) + \mu\frac{Z_{-i}}{Z_i}\right)y^* - \frac{(\omega(L-y)+\mu)\lambda^2\kappa(y^*)^\alpha}{(\lambda-1)(1-\alpha)Z_i}}{\omega^2(L - y^*_\kappa) - \mu^2} \\ &= (\lambda - 1)\frac{\left(\omega(L-y^*) + \mu\frac{Z_{-i}}{Z_i}\right)y^* - \frac{\mu(Z_1+Z_0)y^*}{Z_i}}{\omega^2(L - y^*_\kappa) - \mu^2} \\ &= \frac{(\lambda-1)y^*}{\omega(L-y^*)+\mu} \end{aligned}$$

Hence, $v_i(R^*, R^*)$ is monotonically increasing in y^*, which in turn is monotonically increasing in κ. As noted above, for different values of κ, $y^* \in \left[0, L + \frac{\rho+2\mu}{(1-(1-\alpha)\ln\lambda)\phi}\right)$, which implies v_i ranges between 0 and ∞. The existence of κ^* follows from continuity. ■

Proof of Proposition 4: Consider a fixed value $\kappa > \kappa^*$, where κ^* is defined in Lemma 2. As in the proof of Lemma 2, it will be useful to work with $y_i = L - R_i$. Consider the set

$$S = \{(y_0, y_1) \mid v_0(L-y_0, L-y_1) = v_1(L-y_1, L-y_0)\}$$

The proposition follows if I can show that within this set there exists an element $(y_0, y_1) \in S$ such that $y_0 > y_1$ and $\phi v_i(L-y_i, L-y_{-i}) = 1$ for $i \in \{0,1\}$.

Consider first the case where $y_1 = 0$. For this value, we have

$$\begin{aligned} v_0 &= \frac{(\lambda-1)\omega(L)y_0 - \frac{\omega(L)+\mu\lambda^2\kappa}{(1-\alpha)Z_0}y_0^\alpha}{\omega(L-y_0)\omega(L)-\mu^2} \\ v_1 &= 0 \end{aligned}$$

Hence, there are exactly two values of y_0 for which $v_0(L-y_0, L) = v_1(L, L-y_0)$, namely $y_0 = 0$ and

$$y_0 = \widetilde{y}_0 \equiv \left[\frac{(\omega(L)+\mu)\lambda^2\kappa}{\omega(L)(\lambda-1)(1-\alpha)Z_0}\right]^{\frac{1}{1-\alpha}}$$

By the implicit function theorem, there exist continuous functions $y_0(\cdot)$ defined in a neighborhood of $y_1 = 0$ such that $y_0(y_1) \to 0$ and $y_0(y_1) \to \widetilde{y}_0$ as $y_1 \to 0$ which satisfy $v_0(L - y_0(y_1), L - y_1) = v_0(L - y_1, L - y_0(y_1))$.

For $y_1 \neq 0$, we can rewrite the equation $v_0 = v_1$ in terms of y_1 and $\xi = \frac{y_0}{y_1}$. The condition that $v_1 = v_0$ can be rewritten as

$$\frac{\lambda^2\kappa y_1^{\alpha-1}}{Z_1(\lambda-1)(1-\alpha)}(A_0 - A_1\xi - A_2\xi^\alpha) - 1 + \xi + h(\xi) = 0 \tag{5.5}$$

where

$$\begin{aligned} A_0 &= \frac{\omega(L)+\mu}{\omega(L)} \\ A_1 &= \frac{(1-(1-\alpha)\ln\lambda)\,\phi y_1}{\omega(L)} \\ A_2 &= \frac{\omega(L-y_1)+\mu}{\omega(L)}\frac{Z_1}{Z_0} \end{aligned}$$

and as in Lemma 1,

$$h(\xi) = \frac{\mu}{\omega(L)}\left[\frac{Z_1}{Z_0}\xi^{\alpha} - \frac{Z_0}{Z_1}\xi^{1-\alpha}\right]$$

For notational convenience, I will rewrite (5.5) more compactly as

$$Q(\xi; y_1) = 0$$

The implicit functions $y_0(y_1)$ described above which limit to 0 and $\widetilde{y}_0$ establish the existence of functions $\xi(y_1)$ defined locally near $y_1 = 0$ that limit to $\xi = \left(\frac{Z_0}{Z_1}\right)^{\frac{1}{\alpha}}$ and $\xi = \infty$, respectively, as $y_1 \to 0$.

Define y^* as in Lemma 2. Differentiate $Q(\xi; y_1)$ with respect to ξ twice to obtain

$$\frac{\partial^2 Q}{\partial \xi^2} = h''(\xi) + \alpha(1-\alpha)\frac{\lambda^2\kappa\xi^{\alpha-2}}{Z_1(\lambda-1)(1-\alpha)}A_2 y_1^{\alpha-1}$$

Substituting in for $h''(\xi)$, we obtain

$$\frac{\partial^2 Q}{\partial \xi^2} = \alpha(1-\alpha)\left(\frac{\mu}{\omega(L)}\left[\frac{Z_0}{Z_1}\xi^{-\alpha-1} - \frac{Z_1}{Z_0}\xi^{\alpha-2}\right] + \frac{\lambda^2\kappa\xi^{\alpha-2}}{Z_1(\lambda-1)(1-\alpha)}A_2 y_1^{\alpha-1}\right)$$

Now, note that

$$\begin{aligned} A_2 y_1^{\alpha-1} &= \frac{\omega(L-y_1)+\mu}{\omega(L)}\frac{Z_1}{Z_0}y_1^{\alpha-1} \\ &= \left[\frac{\omega(L)+\mu}{\omega(L)}y_1^{\alpha-1} - \frac{(1-(1-\alpha)\ln\lambda)\,\phi y_1^{\alpha}}{\omega(L)}\right]\frac{Z_1}{Z_0} \end{aligned}$$

is decreasing in y_1. Hence, if we can show that $\frac{\partial^2 Q}{\partial \xi^2} > 0$ for some y_1', it follows that $\frac{\partial^2 Q}{\partial \xi^2} > 0$ for all $y_1 \leq y_1'$. For $y_1 = y^*$, we know from the proof of Lemma 2 that y^* satisfies

$$\frac{\lambda^2\kappa(y^*)^{\alpha-1}}{Z_1(\lambda-1)(1-\alpha)} = \frac{\mu}{\omega(L-y^*)+\mu}\left(\frac{Z_1+Z_0}{Z_1}\right)$$

Substituting this into the expression for $\frac{\partial^2 Q}{\partial \xi^2}$ yields

$$\frac{\partial^2 Q}{\partial \xi^2} = \alpha(1-\alpha)\left(\frac{\mu}{\omega(L)}\frac{Z_0}{Z_1}\xi^{-\alpha-1} + \frac{\mu}{\omega(L)}\xi^{\alpha-2}\right) > 0$$

Hence, for all $y_1 \leq y^*$, $Q(\xi; y_1)$ is convex in ξ. This will prove important in what follows.

Before I proceed, I introduce the notation $(y_0, y_1) \rightsquigarrow (y_0', y_1')$ to denote the case in which there exists a continuous mapping $y_1(\tau) > 0$ and a continuous mapping $y_0(\tau)$ defined for $\tau \in (0,1)$ such that

1. $\lim_{\tau \to 0} y_1(\tau) = y_1$ and $\lim_{\tau \to 0} y_0(\tau) = y_0$
2. $\lim_{\tau \to 1} y_1(\tau) = y_1'$ and $\lim_{\tau \to 1} y_0(\tau) = y_0'$
3. For all $\tau \in (0,1)$, $(y_0(\tau), y_1(\tau)) \in S$, i.e. $Q(\xi(\tau); y_1(\tau)) = 0$ for $\xi(\tau) = \dfrac{y_0(\tau)}{y_1(\tau)}$

The notation $\lim_{\tau \to 1} y(\tau) = \infty$ denotes, as usual, that for every $N > 0$, there exists a τ_N such that $y(\tau) > N$ for all $\tau > \tau_N$. Thus, we can describe a path in which $y_i' = \infty$ for some i. Note that if we can establish that $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$, the statement of the proposition follows from a simple continuity argument: since $\phi v_0(L - y^*, L - y^*) > 1$ for $\kappa > \kappa^*$ but $\phi v_0(\widetilde{y}_0, 0) = 0$, there exists some τ for which $(y_0(\tau), y_1(\tau)) \in S$ and where $\phi v_i(L - y_i(\tau), L - y_{-i}(\tau)) = 1$. Since $v_0(L - y, L - y) = v_1(L - y, L - y)$ if and only if $y = y^*$, and since $\widetilde{y}_0 > 0$, it follows that $y_0(\tau) > y_1(\tau)$ by continuity.

I now break down my analysis into different cases, depending on the sign of $\dfrac{\partial Q}{\partial \xi}$ evaluated at $\xi = 1$ and $y_1 = y^*$. For convenience, a graphical view of the set S corresponding to each of the three cases (for particular parameter values) is provided in Figure A3.

Case I: $\dfrac{\partial Q(1; y^*)}{\partial \xi} > 0$

I claim this is enough to establish that $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$, which from above is enough to establish the proposition. I first argue that there exists a $y_1' \in [0, y^*)$ such that $(y^*, y^*) \rightsquigarrow (y_0', y_1')$ where $\dfrac{y_0'}{y_1'} = \infty$. For suppose not. Since $\dfrac{\partial Q(1; y_1)}{\partial y_1} > 0$ for all y_1, it follows that $Q(1; y_1) < 0$ for all $y_1 < y^*$. Furthermore, since $Q(\xi; y^*)$ is convex in ξ, it also follows that $Q(\xi; y^*) > 0$ for all $\xi > 1$. By continuity, then, the assumption that $(y^*, y^*) \not\rightsquigarrow (\infty, y_1)$ for all $y_1 \in (0, y^*)$ implies that for each $y_1 \in (0, y^*)$ there must exist some $\xi(y_1) > 1$ such that $Q(\xi(y_1); y_1) > 0$. Applying the intermediate value theorem, we can deduce that for every $y_1 \in (0, y^*)$ there exists a $\xi^*(y_1) > 1$ such that $Q(\xi^*(y_1); y_1) = 0$. Since Q is continuous and convex in ξ for all $y_1 \leq y^*$, the root $\xi^*(y_1)$ is the unique value of $\xi > 1$ such that $Q(\xi^*(y_1); y_1) = 1$, is continuous in y_1, and $\lim_{y_1 \uparrow y_\kappa^*} \xi^*(y_1) = 1$. Hence, $(y^*, y^*) \rightsquigarrow (\xi^*(y_1), y_1)$. Taking the limit as $y_1 \downarrow 0$, it follows that $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$, since the unique root greater than one for which $\lim_{y_1 \to 0} Q(\xi; y_1) = 0$ limits to ∞. But then $(y^*, y^*) \rightsquigarrow (y_0', y_1')$ such that $\dfrac{y_0'}{y_1'} = \infty$, which is a contradiction. Since $\lim_{\xi \to \infty} \dfrac{\partial Q}{\partial y_1} < 0$ for all $y_1 > 0$, there can exist at most one y_1 for which $\lim_{\xi \to \infty} Q(\xi; y_1) = 0$. Since $\lim_{\xi \to \infty} Q(\xi; 0) = 0$, it follows that $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$.

Case II: $\dfrac{\partial Q(1; y^*)}{\partial \xi} = 0$.

Since $Q(1; y^*)$ is strictly convex, it follows that $Q(\xi; y^*) > 0$ for all $\xi \neq 1$. The fact that $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$ then follows from the same argument as in Case I.

Case III: $\dfrac{\partial Q(1; y^*)}{\partial \xi} < 0$

In this case, the arguments above can no longer be used to establish that $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$. However, I argue that if $(y^*, y^*) \not\rightsquigarrow (\widetilde{y}_0, 0)$, then there exists some $y_1 > 0$ such that $(y^*, y^*) \rightsquigarrow (\infty, y_1)$ where $\lim_{\tau \to 1} \dfrac{y_0(\tau)}{y_1(\tau)} > 1$ along any such connecting path. As I argue below, this condition is also sufficient to establish the proposition.

Again, the proof is by contradiction. Suppose the claim is false, i.e. suppose $(y^*, y^*) \not\rightsquigarrow (\widetilde{y}_0, 0)$ and $(y^*, y^*) \not\rightsquigarrow (\infty, y_1)$ for all $y_1 > 0$, including $y_1 = \infty$. If we differentiate ξ with respect to y_1 along the curve $Q(\xi; y_1) = 0$, we obtain

$$\left.\frac{d\xi}{dy_1}\right|_{(\xi; y_1) = (1; y^*)} = -\frac{\partial Q / \partial y_1}{\partial Q / \partial \xi} > 0$$

where the last inequality follows from the fact that $\dfrac{\partial Q(1; y_1)}{\partial y_1} > 0$ for all y_1. Hence, if $(y^*, y^*) \rightsquigarrow (y_0, y_1)$ for some $y_0 > y^*$, it follows from continuity and the uniqueness of y^* that $y_0 > y_1$. Next, since $(y^*, y^*) \not\rightsquigarrow (\infty, y_1)$ for all y_1 (including $y_1 = \infty$) by assumption, it follows that

$$\overline{y} = \sup \{ y_0 \mid (y^*, y^*) \rightsquigarrow (y_0, y_1) \text{ for some } y_1 \}$$

is finite. It follows that for any $y_0 > y^*$, it must be the case that $(y^*, y^*) \not\rightsquigarrow (y_0, y_1)$ if $y_1 \geq \overline{y}$. Thus, any continuous path that originates at (y^*, y^*) for which $y_0(\tau) > y^*$ is bounded in its y_1 term from above by $\overline{y}$. But the fact that $\lim_{y_1 \to 0} Q(\xi; y_1) < 0$ for all finite $\xi > 0$, together with continuity and the uniqueness of y^*, implies that this occurs only if $(y^*, y^*) \rightsquigarrow (y_0, y_1)$ for some $y_0 > y^*$ and some y_1 such that $\lim_{\tau \to 1} \dfrac{y_0(\tau)}{y_1(\tau)} = \infty$. Since $y_0(\tau) \leq \overline{y}$ for all τ, this requires that $\lim_{\tau \to 1} y_1(\tau) = 0$. But this contradicts the fact that $(y^*, y^*) \not\rightsquigarrow (\widetilde{y}_0, 0)$. It follows that either $(y^*, y^*) \rightsquigarrow (\widetilde{y}_0, 0)$ or $(y^*, y^*) \rightsquigarrow (\infty, y_1)$ where $\lim_{\tau \to 1} \dfrac{y_0(\tau)}{y_1(\tau)} > 1$ along this path.

The final step is to prove that the fact that $(y^*, y^*) \rightsquigarrow (\infty, y_1)$ where $\lim_{\tau \to 1} \dfrac{y_0(\tau)}{y_1(\tau)} > 1$ implies there exists a solution (y_0, y_1) with $y_0 > y_1$ such that $\phi v_0(L - y_0, L - y_1) = \phi v_1(L - y_1, L - y_0) = 1$. Consider

$$\lim_{y_i \to \infty} v_i(L - y_i, L - y_{-i}) = \lim_{y_i \to \infty} \frac{(\lambda - 1)\left[\omega(L - y_{-i}) y_i + \left(\mu \frac{Z_{-i}}{Z_i} y_{-i}^{1-\alpha} - \frac{(\omega(L - y_{-i}) + \mu)\lambda^2 \kappa}{(1-\alpha)(\lambda - 1) Z_i}\right) y_i^{\alpha}\right]}{\omega(L - y_i)\,\omega(L - y_{-i}) - \mu^2}$$

As $y_i \to \infty$, the numerator converges to $\pm\infty$, depending on the sign of $\omega(L - y_{-i})$, and since $\omega(L - y_i) \to -\infty$, the denominator converges to $\pm\infty$, again depending on the sign of $\omega(L - y_{-i})$. Applying L'Hopital's rule, we obtain

$$\lim_{y_i \to \infty} v_i(L - y_i, L - y_{-i}) = -\frac{(\lambda - 1)\,\omega(L - y_{-i})}{\omega'(\cdot)\,\omega(L - y_{-i})} < 0$$

Hence, since $v_0(L - y_0, L - y_1) < 0$ as $y_0 \to \infty$, it follows by continuity that there exists a pair (y_0, y_1) such that $\phi v_i(L - y_i, L - y_{-i}) = 1$. Again, since $v_0(L - y, L - y) = v_1(L - y, L - y)$ if and only if $y = y^*$ and $\lim_{\tau \to 1} \dfrac{y_0(\tau)}{y_1(\tau)} > 1$, it follows that $y_0(\tau) > y_1(\tau)$ by continuity.

Remark: all cases for $\frac{\partial Q(1;y^*)}{\partial \xi}$ are possible, depending on parameter values. In cases II and III, there will be multiple solutions to the problem $\phi v_i(R_i, R_{-i}) = 1$, i.e. in addition to the solution identified above, there also exists a second solution with $R_0 > R_1$. However, the existence of multiple solutions does not necessarily imply multiple equilibria, since these solutions may involve negative values of R_i. ■

Proof of Proposition 5: Given that firms maximize expected profits, a firm that has successfully innovated will choose the time to implement by solving

$$\max_s E_t \left[e^{-\rho s} \frac{1/P_{t+s}}{1/P_t} v_{t+s} \right]$$

Suppose the solution in Proposition 4 is an equilibrium. The Proposition follows if we can show that $s = 0$. Clearly, if $Z_{t+s} = Z_t$, there is no benefit from delay, since $v_{t+s} = v_t$ and so in the best case scenario the firm becomes the leader and earns v_t discounted at a positive rate. Thus, given other agents are implementing immediately and innovating in accordance with Proposition 4, a firm will only delay implementation until a change in the level of productivity. If the current level of productivity is equal to Z_1, then given $v_0 = v_1 = \frac{1}{\phi}$ in equilibrium, waiting until a regime change yields at most

$$\frac{\mu}{\rho+\mu} \frac{P_1}{P_0} v = \frac{\mu}{\rho+\mu} \frac{Z_0}{Z_1} \left(\frac{L-R_1}{L-R_0} \right)^{\alpha} v$$

assuming the firm is the leading producer when it implements. Since $R_1 > R_0$, it follows that this is less than v. Thus, there is no reason to delay an innovation uncovered when productivity is high. Conversely, there will be no reason to delay an innovation that is discovered when productivity is low if

$$\frac{\mu}{\rho+\mu} \frac{Z_1}{Z_0} \left(\frac{L-R_0}{L-R_1} \right)^{\alpha} < 1$$

By assumption, $\frac{\mu}{\rho+\mu} \frac{Z_1}{Z_0} < 1$. Moreover, by continuity, the solution (R_0, R_1) identified in Proposition 4 limits to (R^*, R^*) as $\kappa \to \kappa^*$. Thus, there will be no benefit from delay even though $R_1 > R_0$ for κ close to κ^*.

References

[1] Aghion, Philippe and Peter Howitt, 1992. "A Model of Growth through Creative Destruction" *Econometrica*, 60(2), March, p323-51.

[2] Aghion, Philippe and Gilles Saint Paul, 1998. "Virtues of Bad Times: Interaction between Productivity Growth and Economic Fluctuations" *Macroeconomic Dynamics*, September, 2(3), p322-44.

[3] Barlevy, Gadi, 2002. "The Cost of Business Cycles under Endogenous Growth" Mimeo, Northwestern University.

[4] Barlevy, Gadi and Dani Tsiddon, 2000. "Wage Inequality and the Business Cycle" Mimeo, Northwestern University.

[5] Basu, Susanto, 1996. "Procyclical Productivity: Increasing Returns or Cyclical Utilization?" *Quarterly Journal of Economics*, 111(3), August, p719-51.

[6] Benhabib, Jess and Roger Farmer, 1994. "Indeterminacy and Increasing Returns" *Journal of Economic Theory*, 63(1), June, p19-41.

[7] Betts, Julian and Laurel McFarland, 1995. "Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments" *Journal of Human Resources*, 30(4), p741-65.

[8] Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, 1993. "Labor Hoarding and the Business Cycle" *Journal of Political Economy*, 101(2), April, p245-73.

[9] Caballero, Ricardo and Mohamad Hammour, 1996. "On the Timing and Efficiency of Creative Destruction" *Quarterly Journal of Economics*, August, 111(3), p805-52.

[10] Caballero, Ricardo and Mohamad Hammour, 1999. "The Costs of Recessions Revisited: a Reverse-Liquidationist View" NBER Working Paper #W7355, September.

[11] Canton, Eric and Harald Uhlig, 1999. "Growth and the Cycle: Creative Destruction versus Entrenchment" *Journal of Economics*, 69(3), p239-66.

[12] Cooper, Russell and John Haltiwanger, 1993. "The Aggregate Implications of Machine Replacement: Theory and Evidence" *American Economic Review*, June, 83(3), p181-186.

[13] Cooper, Russell, John Haltiwanger, and Laura Power, 1999. "Machine Replacement and the Business Cycle: Lumps and Bumps" *American Economic Review*, 89(4), September, p921-46.

[14] Davis, Steven and John Haltiwanger, 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation" *Quarterly Journal of Economics*, August, 107(3), p819-63.

[15] Dellas, Harris and Plutarchos Sakellaris, 1997. "On the Cyclicality of Schooling: Theory and Evidence" Mimeo, University of Maryland.

[16] DeJong, David and Beth Ingram, 2001. "The Cyclical Behavior of Skill Acquisition" *Review of Economic Dynamics*, July 2001, 4(3), p536-61.

[17] Fatas, Antonio, 2000. "Do Business Cycles Cast Long Shadows? Short-Run Persistence and Economic Growth" *Journal of Economic Growth*, 5(2), June, p147-62.

[18] Francois, Patrick and Huw Lloyd-Ellis, 2002. "Animal Spirits Meets Creative Destruction" Mimeo, Queen's University and forthcoming, *American Economic Review*.

[19] Francois, Patrick and Shouyong Shi, 1999. “Innovation, Growth, and Welfare-Improving Cycles” *Journal of Economic Theory*, 85, p226-57.

[20] Gomes, Joao, Jeremy Greenwood, and Sergio Rebelo, 2001. “Equilibrium Unemployment” *Journal of Monetary Economics*, 48(1), p109-52.

[21] Griliches, Zvi, 1984. “Comment on R&D and Innovation: Some Empirical Findings” in *R&D, Patents, and Productivity*, edited by Zvi Griliches, Chicago: University of Chicago Press, p148-149

[22] Griliches, Zvi, 1990. “Patent Statistics as Economic Indicators: a Survey” *Journal of Economic Literature*, December, 28(4), p1661-1707.

[23] Grossman, Gene and Elhanan Helpman, 1991. “Quality Ladders in the Theory of Growth” *Review of Economic Studies*, 58(1), January, p43-61.

[24] Hall, Robert 1991. “Labor Demand, Labor Supply, and Employment Volatility” *NBER Macroeconomics Annual.* Cambridge, Mass: MIT Press.

[25] Hall, Robert, 2000. “Reorganization” *Carnegie Rochester Conference Series on Public Policy*, 52(0), June, p1-22.

[26] Jovanovic, Boyan, 2002. “Asymmetric Cycles” Mimeo, NYU and University of Chicago.

[27] Lucas, Robert, 1978. “Asset Prices in an Exchange Economy” *Econometrica*, 46(6), November, p1429-45.

[28] Lucas, Robert, 1987. *Models of Business Cycles.* Oxford: Basil Blackwell.

[29] Mansfield, Edwin, 1987. “Price Indexes for R and D Inputs, 1969-1983” *Management Science*, 33(1), January, p124-9.

[30] Mortensen Dale and Christopher Pissarides, 1994. “Job Creation and Job Destruction in the Theory of Unemployment” *Review of Economic Studies*, July, 61(3), p397-415.

[31] Ramey, Valerie, 1991. “Comment on Markups and the Business Cycle” *NBER Macroeconomics Annual*, vol 6, p134-9.

[32] Rotemberg, Julio and Michael Woodford, 1999. “The Cyclical Behavior of Prices and Costs” in *Handbook of Macroeconomics* (1B), vol. 15, edited by John Taylor and Michael Woodford. Amsterdam: Elsevier Science, p1051-1135.

[33] Shleifer, Andrei, 1986. “Implementation Cycles” *Journal of Political Economy*, 94(6), December, p1163-90.

[34] Young, Alwyn, 1998. “Growth without Scale Effects” *Journal of Political Economy*, 106(1), February, p41-63.

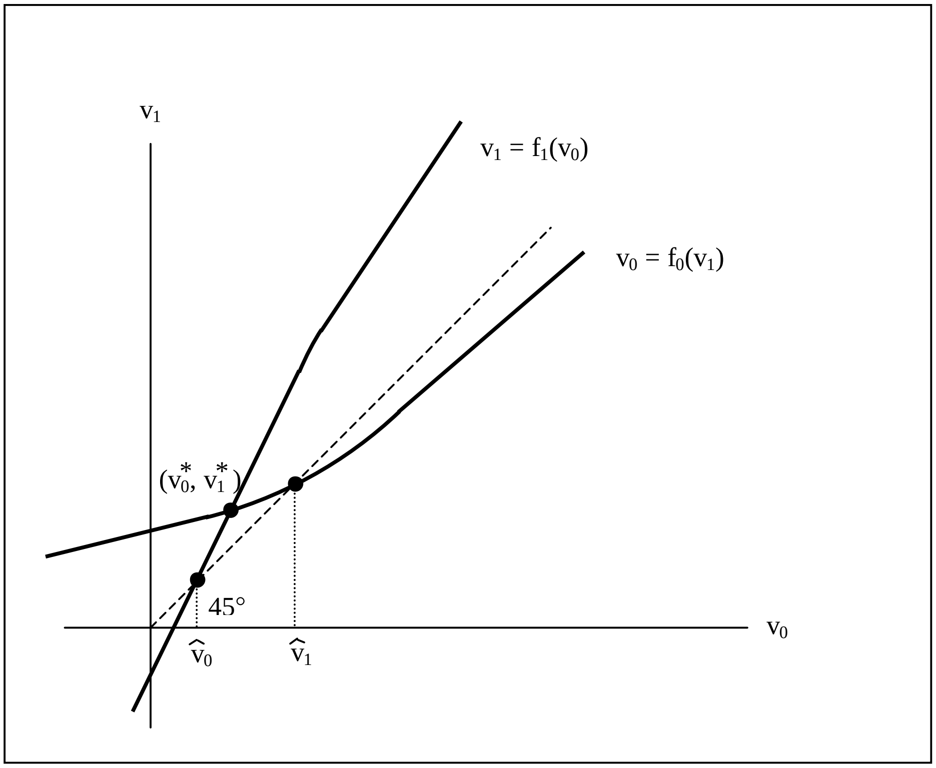


Figure A1: Uniqueness of Social Optimum

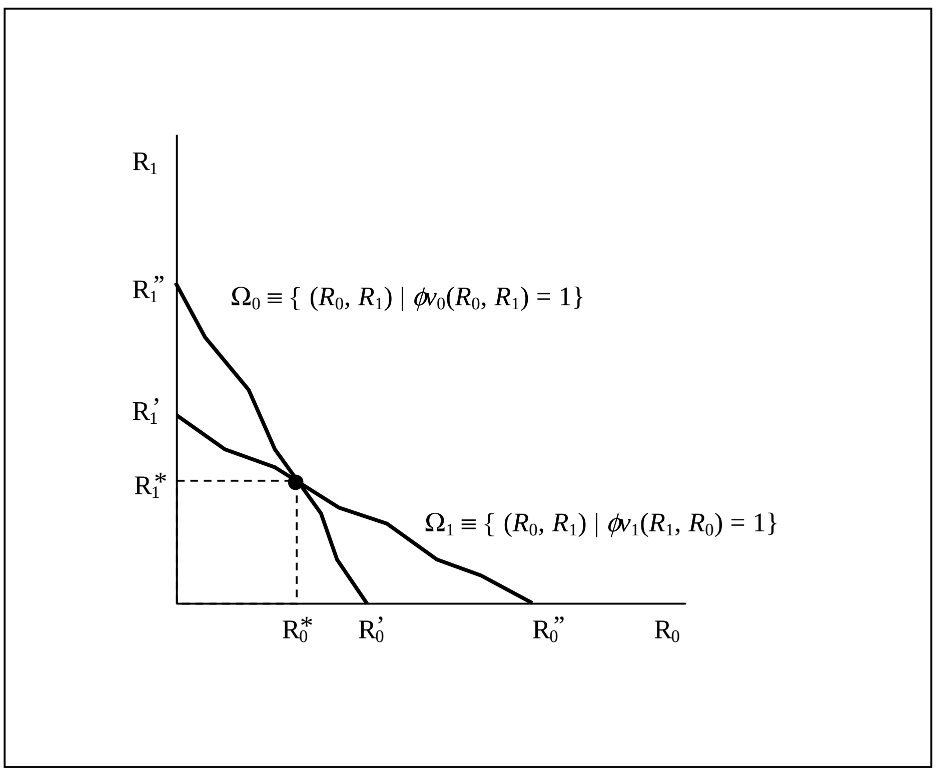
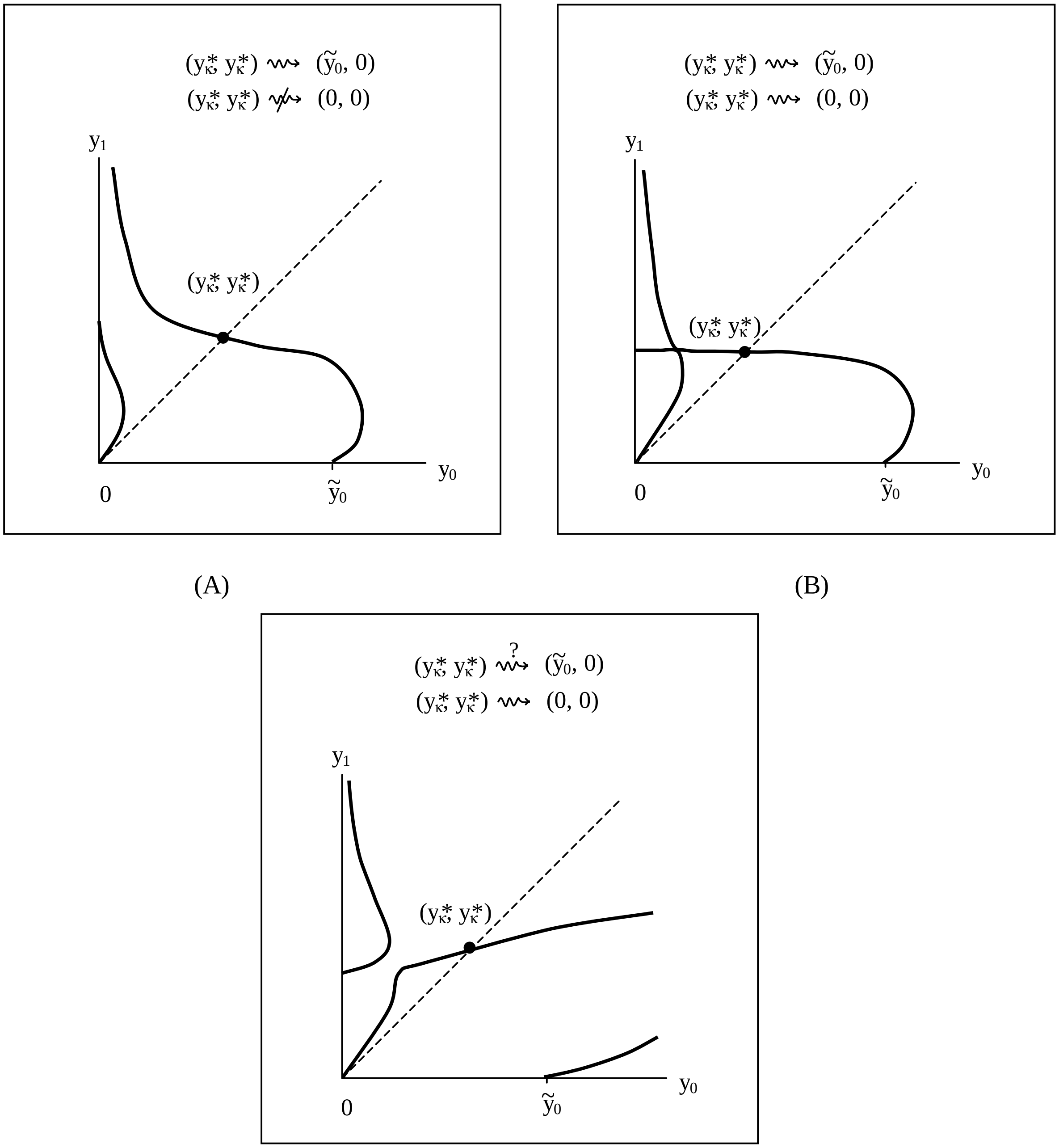


Figure A2: Uniqueness of Markov-Perfect Equilibrium



(C)

Figure A3: The Evolution of the set S in Proposition 4

Panel (A) corresponds to Case I in proof of Proposition 4
Panel (B) corresponds to Case II in proof of Proposition 4
Panel (C) corresponds to Case III in proof of Proposition 4