Innovate to lead or innovate to prevail: When do monopolistic rents induce growth?

Roberto Piazza · Yu Zheng

Received: 31 October 2021 / Accepted: 5 December 2022 / Published online: 8 March 2023
© International Monetary Fund 2023

Abstract
This paper extends the standard Schumpeterian model of creative destruction by allowing the cost of innovation for followers to increase in their technological distance from the leader. This assumption is motivated by the observation that the more technologically advanced the leader is, the harder it is for a follower to leapfrog without incurring extra cost for using leader’s patented knowledge. Under this R&D cost structure, leaders have an incentive to play an “endpoint strategy”: they increase their technological advantage, counting on the fact that followers will eventually stop innovating—allowing leadership to prevail. We find that several results in the standard model fail to hold. In addition to the high-growth steady state in which only followers innovate, there now exists a second saddle-path-stable steady state: a low-growth steady state that features both leaders and followers innovating. A policy that increases monopolistic rents or extends parent duration can push the economy toward the low-growth steady state, causing, in some cases, irreversible harm to long-run growth.

Keywords Innovation · Persistent monopoly · Endogenous growth · Growth trap

JEL Classification O31 · O34 · O41 · L16

Yu Zheng
yu.zheng@qmul.ac.uk

Roberto Piazza
rpiazza2@imf.org

1 Research Department, International Monetary Fund, 700 19th Street, N.W., Washington, DC 20431, USA
2 School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK
3 The Centre for Economic Policy Research (CEPR), London, UK
1 Introduction

The Schumpeterian idea of creative destruction lies at the heart of the innovation-based growth theory, in which technological progress is the outcome of an innovation game where firms undertake costly R&D to invent a better product and leapfrog the industry leaders (Romer 1990; Grossman and Helpman 1991; Aghion and Howitt 1992). There are two key ingredients to this theory of “creative destruction” of leadership. First, a successful innovator (leader), by launching a new product in the market, reveals to potential competitors (followers) the frontier knowledge embodied in the product. Such full knowledge spillover ensures a level playing field in the subsequent round of innovation race for the next product. Second, to prevent this very knowledge spillover from enabling imitators to drive the market price of the new product down to the marginal cost, monopoly rights must be granted to the leader whereby he can recoup his initial investment in R&D. The rich dynamics of competition, firm exit and turnover inherent in the Schumpeterian model makes it the bedrock on which a large and growing literature is based, where the outcomes from creative destruction among heterogeneous firms can be mapped to micro data (e.g. the seminal work of Klette and Kortum 2004).

In this paper, we call into question a key assumption of this Schumpeterian model of leapfrogging innovation and show how some of its main implications get turned on their heads. The assumption is that large knowledge spillovers, made possible for example through the patent’s registration process, enables all, leaders and followers alike, to compete equally in the innovation race for the next product. In reality, this assumption is often violated. Because of the very nature of the patent system, any innovation that builds upon previously patented knowledge would face costly legal challenges by the industry leader. This discourages the use of existing patents by followers who attempt to displace the current industry leader (Boldrin and Levine 2013). And in any case, as any student knows, even if litigation is not a concern, knowledge is never really for free and learning is a costly, time consuming process.

Examples of the disadvantage that a follower faces trying to improve on a leader’s product abound in history and at present. In 1769, the great inventor James Watt obtained a patent on his idea of a separate condenser in a steam engine, an improvement upon the Newcomen steam engine. Over the following thirty years while the patent lasted, steam engines were modified and improved by many of his peers: William Bull, Richard Trevithick, Arthur Woolf, and Jonathan Hornblower. Yet none of these models made it to the market until 1804 after the patent expired, because, no matter how much better the newer models were, they had to use the idea of the separate condenser (Boldrin and Levine 2008 contains many more examples). Fast forward 200 years, in 2007, Apple and Samsung began their decade-long multi-million dollar patent war that spread across courts in ten countries around the globe. This is yet another example of the unequal footing from which a leader (then Apple) and a follower (then Samsung) were competing to bring about new products.

The discussion above suggests that innovation costs for a follower are at least as large as, if not larger than, those for a leader. Moreover, the technologically more advanced a leader is relative to the follower, the more costly it is for the follower to leapfrog him. At the extreme, when the technological distance is large enough,
then the leader has achieved the “endpoint strategy” (Hörner 2004) of pushing the innovation race to a state where any attempt by the follower to leapfrog the incumbent has become prohibitively expensive, and innovation efforts by both the leader and the follower cease. The asymmetry between leader’s and follower’s innovation capability has long been addressed by the theoretical microeconomics literature on races (tracing back to Harris and Vickers 1987; Budd et al. 1993, and Hörner 2004). It is exactly the intuition developed in this literature that we add to our endogenous growth model. To put it simply, we embed the assumption of state-dependent asymmetric R&D costs into an otherwise standard Scumpeterian model à la Grossman and Helpman (1991) (henceforth, GH), where followers’ R&D efforts are aimed at leapfrogging the leaders and thus contribute to aggregate growth.

More formally, we study the effects of state-dependent innovation costs in a general equilibrium model with a continuum of industries, where in each industry a leader and a competitive fringe of followers (which we refer to simply as “followers”) play a game of innovation. The state of the industry is the technological distance between the leader and the followers, such that when followers fall behind in the innovation race they see their innovation costs rise. We show that, in this model, the balance growth path where only followers innovate (as in GH) is no longer the only one that can emerge in equilibrium. For a certain range of parameters, the equilibrium of the GH type, which we denote with $H$ (standing for “high-growth”), can coexist with growth traps, which are low-growth equilibria (denoted by $L$) where also leaders innovate provided that their technological gap with the followers is small. These equilibria are characterized by low growth because the R&D effort by leaders has two opposite effects on the aggregate innovation rate. On one hand, it contributes to raising the innovation rate of an industry (the intensive margin of innovation) when the leader and its followers are close to each other. On the other, it increases the share of industries (the extensive margin) where leaders and followers are sufficiently far from each other that innovation drops to zero. In our model the positive effect on the intensive margin is always dominated by the negative effect on the extensive margin, so that the success of leaders’ end-point strategies of kicking the followers out of the innovation race is both the reason for large innovation being done at times by leaders, and the cause of low aggregate growth.

The structure of the steady states enriches the growth implications from policies that regulate market power. In the GH setting, increasing monopoly power only provides incentives for followers to innovate, and therefore always raises steady state growth. In contrast, in our setting, the positive link between monopoly power and long-run steady state growth is broken: under certain conditions, even a small increase in monopoly power can result in large reduction in aggregate steady state growth. The mechanism behind this result is intrinsically related to the structure of the steady states. When the economy is initially in an $H$ steady state, an increase in monopoly protection through, for example, an increase in the monopolist’s markup ceiling or a lengthening of patent duration, can tip the economy into the parameter region where only the $L$ steady state exists. As the economy begin transitioning toward the $L$ steady state, it initially experiences a short-term bout of growth, caused by leaders’ aggressive R&D investment. Over time, as the share of zero-innovation industries where leaders and followers are technologically far apart increases, aggregate growth falls and eventually
settles down to its $L$ value. This mechanism speaks in spirit to recent empirical evidence of increasing market power, declining investment and business dynamism in the US (Haltiwanger 2015; Covarrubias et al. 2019; Loecker et al. 2020). As an additional result, the paper also shows that the existence of multiple steady states poses a “non-reversibility” conundrum to the policy maker. In fact, should the policy maker wish to reverse course and reduce the degree of monopolist protection, he may find that simply undoing his original policy change is not enough to restore the initial $H$ steady state, and that a much larger policy correction is needed.

There have been many papers in the literature that feature leaders innovating and persistent monopoly. Some of the papers emphasize the asymmetry of the technology of innovation between leaders and followers but without the technology being state-dependent (Barro and Sala-i-Martin 1995; Segerstrom and Zolnierek 1999; Segerstrom 2007). Some focus on the asymmetry between leaders and followers in aspects of production other than innovation: customer’s base (Stein 1997) and overhead or expansion cost (Klette and Kortum 2004; Aghion et al. 2022). Others examine asymmetry stemming from the particular market or game structure that leaders and followers are in, for example a leader’s first mover advantage or initial market power (Gilbert and Newbery 1982; Denicolo 2001; Etro 2004; Aghion et al. 2005). Among the aforementioned papers, several share our message that persistent monopolies can be detrimental to long-run growth, though it’s clear that we reach the conclusion from very different mechanisms (Gilbert and Newbery 1982; Aghion et al. 2005, 2022).¹

Following the tradition of the classic Schumpeterian framework (Romer 1990; Grossman and Helpman 1991; Aghion and Howitt 1992), we adopt a setup where a competitive fringe innovates to leapfrog the industry leader. A real world example is highly innovative start-ups attempting to disrupt an industry and replace its incumbent. This is in contrast and complementary to a related but different strand of literature which models the interaction between the industry leader and follower as a duopoly and which is known as the “step-by-step” model of innovation (Aghion et al. 2001).² Recent work along this line includes Acemoglu and Akcigit (2012) and Liu et al. (2022). In those models, due to the “escape competition” effect, leaders also have incentive to innovate and their incentive depends on the relative market shares of the duopolists

¹ Some recent works investigate the impact of patent protection on aggregate growth in models that, different from ours, feature horizontal innovation and heterogeneous households (Chu et al. 2021; Kiedaisch 2021). In a similar vein, Atkeson and Burstein (2019) conduct a quantitative analysis based on the model in Garcia-Macia et al. (2019), which allows leaders to have different innovation rates than followers both vertically and horizontally, to explore the effect of innovation subsidies. Perhaps more similar to us, Suzuki (2021) shows that, in a GH setting but with Cournot competition, cutting the corporate income tax can also reduce growth.

² There are two assumptions that differentiate “step-by-step” models from the traditional Schumpeterian framework to which our paper belongs. First, followers in “step-by-step” models are duopolistic “insiders” who earn non-zero profits. Correspondingly, their R&D technology features strictly increasing marginal costs of innovation. Second, before gaining the lead in the innovation game, followers need to catch up with the industry’s incumbent and reach a “neck-and-neck” state where both firms find themselves in a low-profit situation which they have a strong incentive to escape. In traditional Schumpeterian models, on the contrary, innovation is done by outside followers who, because of the assumption of free entry, earn zero profits until they become the new monopolist. The free-entry assumption justifies the use of linear R&D technologies for followers. In addition, there is no “neck-and-neck” state and product market’s profits of the leader are independent of the distance from the followers.
in the product market. It is well-known in that literature that some degree of product market competition and imitation is growth enhancing. In our Schumpeterian setting, instead, at each instant in time and in each industry there is only one firm (the leader) who produces the final good. The leader’s market share is always equal to one and his profit doesn’t even depend on his distance from the followers. Therefore leaders have no incentive to innovate just for the sake of increasing their market profit (the so-called “Arrow” effect in Schumpeterian models). Instead, our leader’s incentive to innovate is purely based on the goal of increasing followers’ leapfrogging costs, thus deterring followers’ innovation. By studying the interaction of state dependent innovation costs in a simple Schumpeterian model of growth, we are able to characterize the equilibrium analytically, uncovering the characteristics of the multiple steady states which are key to our policy implications.

The rest of the paper proceeds as follow. Section 2 introduces the baseline model, Sect. 3 characterizes the steady states of the economy, while Sect. 4 discusses its global dynamics properties. In Sect. 5, we discuss the implications of our model for policy and welfare. Conclusions follow.

2 The model

The model is based on GH’s seminal model of quality ladders. It is a continuous time infinite horizon model. There is a continuum of goods, indexed by the real numbers in a unit interval. There are two types of agents in the model, households and firms.

2.1 Households

There is a representative household who decides what to consume at each point in time, given its income. It is endowed with one unit of labor and supplies it inelastically. It owns the firms in the economy and hence receives a stream of profits from them. Its wealth at time 0, \( W_0 \), is the present value of the stream of profits and labor income it receives ad infinitum. At each instant, the household chooses the quantity, \( d_{it} \), of each of the \( i \in [0, 1] \) goods to consume, taking as given their quality \( q_{it} \), their price \( p_{it} \), and the instantaneous interest rate \( r_t \).

The household consumes \( C_t \) at time \( t \), which is an aggregate of all varieties of goods:

\[
\log C_t = \int_{[0,1]} \log (q_{it} d_{it}) \, di. \tag{1}
\]

The functions \( q_{it} > 0 \) define the highest quality developed up to time \( t \) for good \( i \). The household’s lifetime utility is characterized by a time-additive log period utility function with a rate of time preference of \( \rho \). The household’s problem is:

\footnote{Assuming log period utility simplifies the algebra, but all main results are robust to assuming instead a more general utility function with constant intertemporal elasticity of substitution (Appendix E).}
\begin{align*}
\max_{\{d_{it}, \forall i\}_{t=0}} & \int_0^\infty e^{-\rho t} \log C_i dt \\
\text{s.t.} & \int_0^\infty e^{-R_t} E_t dt \leq W_0,
\end{align*}

where $R_t$ is the compounded interest rate and $E_t$ represents total spending at time $t$:

$$
R_t = \int_0^t r(t')dt',
$$

$$
E_t = \int_{[0,1]} p_{it} d_{it} dt.
$$

The Cobb-Douglas form of the consumption aggregate implies that the amount spent by the household on good $i$ is the same across all products, giving

$$
d_{it} = \frac{E_t}{p_{it}}.
$$

The intertemporal Euler equation gives

$$
\frac{\dot{E}_t}{E_t} + \rho = r_t.
$$

Household’s wealth $W_0$ is given by

$$
W_0 = \int_0^\infty e^{-R_t} \left[ \Pi_t + w_t L_t + w_t \int_{i \in [0,1]} \omega_{it} \Lambda_{it} di \right] dt,
$$

where $\Pi_t$ are aggregate profits received from firms, $w_t$ is the wage paid to labor employed in the production sector, $L_t$, and $\omega_{it}$ is the wage premium paid to (skilled) labor employed in the R&D sector in industry $i$, $\Lambda_{it}$. We refer to $\Lambda_t$ as the intensive margin of innovation in industry $i$. The role of these variables is explained in detail in Sect. 2.2, which lays out the firms’ problems. Here we simply specify that total labor is in fixed supply, normalized to unity

$$
L_t + \int_{i \in [0,1]} \Lambda_{it} di = 1,
$$

where $L_t$ and $\Lambda_{it}$ are all non-negative. We also assume that the intensive margin of innovation must be bounded above by some constant $\Lambda$. The interpretation is simply that there is at most an amount $\Lambda$ of workers in the economy with the necessary skill to perform R&D activities in any given industry. For example, there is a maximum supply of labor skilled in biomedical sciences available to the pharmaceutical industry, a maximum supply of labor skilled in computer science available to the information technology industry, and so on and so forth. Clearly, in this situation the household’s optimal supply of skilled labor in any given industry corresponds to the entire interval.
Innovate to lead or innovate to prevail...

\[ [0, \Lambda] \text{ if the wage premium is equal to one, while it equals to } \Lambda_i^* = \Lambda \text{ whenever } \omega_i > 1. \]

Modeling the supply of skilled labor as perfectly elastic up to \( \Lambda \) and perfectly inelastic afterwards has two advantages. First, when \( \Lambda \) does not bind, our model is equivalent to GH’s model, so that our model nests GH as a special case. Second, this is a simple and intuitive way to introduce decreasing returns to innovation at the industry level as the cost of skilled labor becoming increasingly more expensive when an industry’s R&D reaches the \( \Lambda \) threshold. We take \( \Lambda \) as an arbitrarily large constant.\(^4\)

### 2.2 Firms

Each product \( i \) corresponds to an industry. In each industry, there is a leader and a competitive fringe of followers. The leader in industry \( i \) has the technology to produce the state-of-the-art version \( q_{it} \) of product \( i \). Such technology is protected by a patent, so that only the leader can produce the quality \( q_{it} \). Leaders and followers also carry out R&D activities. A successful innovation by either a leader or a follower raises the state-of-the-art quality from \( q_{it} \) to \( \gamma q_{it} \), where \( \gamma > 0 \) is the distance between two consecutive rungs on product \( i \)’s quality ladder. The quality of a good can then be written as \( q_{it} = \gamma^s_{it} \), with \( s_{it} \in \mathbb{N} \) the number of rungs along the quality ladder that have been climbed up to time \( t \) in industry \( i \). A detailed description of firms’ production and R&D activities follows.

#### 2.2.1 Production of goods

The output \( y_{it} \) in industry \( i \) is produced using labor \( l_{it} \) according to a linear technology

\[
y_{it} = l_{it}.
\]

Since, for a given industry, products of different qualities are perfect substitutes, a leader who charges a markup over marginal cost of production labor anywhere between 1 and \( \gamma \) can put his followers out of business. Let the markup charged by leaders be \( m \in [1, \gamma] \), which we interpret as a policy variable exogenously determined, as when, for instance, an antitrust authority limits the monopoly pricing power of the leaders. We will later investigate how the growth rate of the economy varies with the markup level \( m \). The price at which leaders sell their products is therefore given by

\[
p_{it} = mw_t, \text{ for } m \in [1, \gamma],
\]

where \( w_t \) is the wage rate of production labor. Without loss of generality, we normalize \( w_t \) to 1 so that, from now on, we express variables in terms of the period wage. The goods prices are then \( p_{it} = m \) and profits of leaders can then be simply expressed as

\[ p_{it} = m w_t, \text{ for } m \in [1, \gamma], \]

\[ (5) \]

\[ 4 \text{ On a technical note, to obtain an equilibrium under a discontinuous labor supply correspondence, we proceed in two steps. First, we propose a continuously differentiable labor supply function with a parameter that governs the speed at which the supply increases as } \lambda \text{ exceeds } \Lambda. \text{ Second, we let the parameter go to infinity to obtain the formulation of the } \Lambda \text{ described above. Details can be found in Appendix A.1.} \]
\[ \pi_{it} = (m - 1) y_{it}. \]

At equilibrium prices, the household’s demand for good \( i \) is given by
\[ d_{it}^* = \frac{E_t^*}{m}. \]

Using the market clearing condition, \( d_{it}^* = y_t^*(i) \), we conclude that all industries produce the same amount of output, \( Y_t \), using the same amount of labor, \( L_t \), given by
\[ Y_t^* = L_t^* = \frac{E_t^*}{m}. \] (6)

It follows then that, for all leaders, profits are given by
\[ \Pi_t^* = (m - 1)L_t^*. \] (7)

2.2.2 Game of innovation

Within each industry, leaders and followers play a game of innovation, and expectations about each other’s future strategies determine current innovation efforts. We start by describing the innovation technologies available to the leaders and followers, which depend on the existing technological distance between the two parties, measured in number of steps on the quality ladder.

When the distance between a leader and a follower is one step, we maintain the GH assumption that both the leader and follower have the same R&D technology. That is, if a firm hires an amount \( \lambda \) of skilled workers to perform R&D, the firm experiences an arrival of a successful innovation at a Poisson rate \( \Gamma(\lambda) \) given by
\[ \Gamma(\lambda) = \chi \lambda, \text{ for } \chi > 0, \]
where \( \chi \) is a parameter that governs the productivity in the R&D sector. The innovation technology displays constant returns at the firm’s level and thus, as we shall see, ensures that the value of followers from entering the innovation game is zero.

When the technological distance between a leader and a follower is instead two or more steps, the follower can no longer innovate with the same technology as that used by the leader. Specifically, we assume that the cost of innovation for the follower who is two or more steps behind the leader is high enough that the follower stops innovating completely. In Appendix B, we show that this assumption is without loss of generality, because, under the assumption of linear innovation technologies, followers who are two or more steps behind the leader would always optimally choose not to innovate.\(^5\)

\(^5\) The assumption of linear technology is not essential for our results. In Appendix F, we solve numerically the model with a quadratic cost function of innovation and show that the main results carry through. However, once we deviate from the linearity assumption and embrace a duopolistic setting with strictly increasing marginal costs, the property that followers stop innovating when they are two-steps behind needs to be assumed and is thus not an equilibrium outcome as in the case with linear technologies (Appendix B).
This structure of the innovation technology is meant to capture the idea that leapfrogging becomes increasingly difficult for followers when their technological distance from the industry leader increases. There are two complementary interpretations for this assumption.

The first is that every state-of-the-art version of a product incorporates elements from the previous versions, which are patented. If the follower’s technology is not far from that of the leader (i.e. when the follower is only one step behind), then the follower is able to invent a new product quality without having to incorporate in this new quality any technological element over which only the leader owns a patent. Indeed, patents can impose substantial legal and uncertainty costs for challengers. Therefore, when followers own patents on quite obsolete technologies (i.e. when the follower is two steps behind the leader), then it is not possible for them to invent the state-of-the-art quality without having to incorporate elements that have already been patented by the leader. As in GH, leaders do not have any incentive to grant a license to a follower, whose innovation costs therefore become prohibitively large.

The second is that some free-of-charge knowledge spillovers to followers do take place, but it takes time. If a leader is a lot more advanced in his stock of knowledge, then it takes a longer time for the knowledge spillovers to complete. In this case, followers fall for some time behind the leader in the amount of R&D knowledge they can muster when the distance to the leader is larger.

When leaders and followers are one step apart, there are potential incentives for both to innovate. Followers innovate to replace the leader, as in GH. Leaders may also want to innovate for the pure goal of distancing themselves further from the followers. As the distance grows, innovation costs for followers rise and the followers stop threatening the incumbent’s leadership, which is then secured. We refer to this strategy of the leader as an endpoint strategy.

We assume that when a leader is two steps ahead of a follower, the distance is reduced to one step at an exogenous (small) rate $\tau > 0$. When rising R&D costs are interpreted as driven by legal constraints imposed by patents, then $\tau$ can be thought as a policy variable that controls the legal term of patents. When rising R&D costs are tied to lack of full knowledge spillover, then $\tau$ indicates the frequency at which the spillover occurs. It is worth noting that when we let $\tau$ go to infinity, we are back to the GH model formulation where there is instant spillover of the innovation technology and leaders and followers can be at most one step apart.

We say that an industry is in the contestable state if the distance between a leader and his followers is equal to one step, and in the non-contestable state if the distance is two steps. We indicate with $\alpha_t \in [0, 1]$ the share of industries that at a given time $t$ are in the contestable state. Since innovation only takes place in a contestable state, we call $\alpha_t$ the extensive margin of innovation in the economy.

Mathematically, the combination of two types of firms (leader $l$ or follower $f$) and of two possible distances (1 or 2) between firms, gives rise to four value functions $V_{\Delta}^j(t)$, for $j \in \{l, f\}$ and $\Delta \in \{1, 2\}$. When it does not create confusion, we omit the indication of the dependence of variables on time. Our four value functions satisfy, at any point of differentiability, the following differential equations
\[ r V^l_2 = \Pi + \tau (V^l_1 - V^l_2) + \dot{V}^l_2 \]  
\[ r V^f_2 = \tau (V^f_1 - V^f_2) + \dot{V}^f_2 \]  
\[ r V^l_1 = \max_{\lambda^l \geq 0} (\pi - \rho \lambda^l + \chi \lambda^l (V^l_2 - V^l_1) + \chi \lambda^l (V^l_1 - V^l_1) + \dot{V}^l_1) \]  
\[ r V^f_1 = \max_{\lambda^f \geq 0} (-\rho \lambda^f + \chi \lambda^f (V^l_1 - V^f_1) + \chi \lambda^f (V^f_2 - V^f_1) + \dot{V}^f_1). \]

Equation (8) (Eq. (9)) describes the value function of a leader (follower) who is two steps ahead (behind). This corresponds to a non-contestable state where thus both the leader and the follower optimally decide not to innovate. Note that the leader is the only one who makes profits and that the industry distance is subject to the exogenous rate \( \tau \) of reduction back to one step. Equations (10) and (11) describe the value functions in a contestable state, where strictly positive innovation rates may still be chosen by both the leader and follower. The leader pays the cost of innovation, \( \omega \lambda^l \), to increase the probability of enlarging the technological gap and obtaining \( V^l_2 \), whereas the follower pays the cost \( \omega \lambda^f \) to increase the probability of leapfrogging and obtaining \( V^f_1 \). The free entry condition for the competitive fringe of followers implies that
\[ V^f_1 (t) = V^f_2 (t) = 0, \quad \forall t. \]

2.3 Equilibrium

In equilibrium, R&D strategies are symmetric across all industries. We focus on Markov equilibria. Therefore, at any given point in time, efforts \( \lambda^l_i \) and \( \lambda^f_i \) by leaders and followers, and the corresponding intensive margin \( \Lambda_t \), are the same across all industries in the contestable state, so we can drop the index \( i \) from our notation. The evolution of the extensive margin is
\[ \dot{\alpha}_t = (1 - \alpha_t) \tau - \alpha_t \frac{\lambda^l_t}{\Lambda_t}. \]

The aggregate number of rungs on the quality ladder achieved at time \( t \), \( S_t = \int_{[0,1]} s_i d i \), evolves according to
\[ \dot{S}_t = \chi H_t \equiv \chi \Lambda_t \alpha_t, \]
where \( H_t \) is defined to be the total amount of skilled R&D labor employed at time \( t \).

The definition of equilibrium in this model is standard.

**Definition 1** An equilibrium is given by prices \( \{r_t, w_t, p_{it}, \omega_t\}_{t=0}^{\infty} \), innovation rates by leaders and followers \( \{\lambda^l_t, \lambda^f_t\}_{t=0}^{\infty} \), functions \( \{E^*_t, L^*_t, \Lambda^*_t, S^*_t, Y^*_t, \Pi^*_t, \alpha^*_t\}_{t=0}^{\infty} \) for aggregate expenditure, supply of production labor, supply of R&D labor, aggregate quality, output, profits and the extensive margin of innovation, such that
i) Given prices and the evolution of \( \Pi^*_t \), the innovation rates \( \lambda^l_t \) and \( \lambda^f_t \) solve firms’ innovation game.
ii) Given aggregate expenditure \( E_t^* \) and normalized wages \( w_t \) are the optimal output and price level chosen by leaders in any industry. Correspondingly, \( \Pi_t^* = (m - 1)Y_t^* \) are the profits of leaders.

iii) Given prices and the evolution of aggregate profits, \( E_t^* \), \( L_t^* \) and \( \Lambda_t^* \) are, respectively, the optimal expenditure, and the optimal production and R&D labor supplies of households.

iv) Given \( \Lambda_t^* \), the wage premium \( \omega_t \) of firms in the contestable state is compatible with fixed supply of skilled labor and satisfies the complementary slackness condition:

\[
(\omega_t - w_t)(\Lambda_t - \Lambda_t^*) = 0.
\]

Given initial condition \( S_0 \) and the evolution of \( H_t^* = \alpha_t^* \Lambda_t^* \), the aggregate quality \( S_t^* \) satisfies (13).

v) Markets clear, i.e. \( L_t^* = Y_t^* \), \( E_t^*/m = Y_t^* \), \( \Lambda_t^* = \lambda_f^* + \lambda_l^* \), \( H_t^* = 1 - L_t^* \).

Note that, since in equilibrium the quantities of all goods are the same and equal to the production labor input, \( d_{it} = L_t \), log aggregate consumption can be written as

\[
\log C_t = \int_{[0,1]} \log(q_{it}d_{it})di = \int_{[0,1]} \log \gamma^{s_{it}}di + \log L_t = \log(\gamma)S_t + \log L_t.
\]

The growth rate of consumption is therefore

\[
\frac{\dot{C}_t}{C_t} = \log(\gamma)\dot{S}_t + \frac{\dot{L}_t}{L_t} = \log(\gamma)\chi H_t + \frac{\dot{L}_t}{L_t} = \log(\gamma)\chi \alpha_t \Lambda_t + \frac{\dot{L}_t}{L_t}.
\]

The growth rate of aggregate consumption is then given, in equilibrium, by the sum of the growth rate of the production labor input and the growth of the aggregate output quality. We refer to the quantity \( g_t = \log(\gamma)\dot{S} \) as the rate of technological growth.

The balanced growth path of this model is an equilibrium where aggregate consumption and quality, \( C_t \) and \( S_t \), grow at the same rate \( g \).

3 Steady states

Depending on the parameter values, the equilibrium economy can display up to three steady states, which we label using subscripts \( H, M \) or \( L \) to indicate whether a steady state is characterized by a high, medium or low value for the extensive margin of innovation, \( \alpha \).

3.1 The \( H \) steady state

The highest possible steady state value for \( \alpha^* \) is 1. In this steady state only followers innovate, and thus \( \lambda_l^* = 0 \) and \( \lambda_f^* > 0 \). As already discussed, by taking \( \overline{\Lambda} \) large enough we can make sure that in a neighborhood of the steady state \( \Lambda_t^* = \lambda_f^* < \overline{\Lambda} \), giving the skill premium \( \omega_t \) equal to 1. Hence, the first order condition for \( \lambda_f \) in a neighborhood of a \( H \) steady state imply that
\[ V_l^i(t) = \frac{1}{\chi}. \]

The condition above naturally implies that \( \dot{V}_l^i = 0 \). Since \( \lambda_l^i = 0 \), a straightforward substitution in the definition of \( V_l^i \) gives

\[ \frac{r_l}{\chi} + \lambda_l^i = \Pi_l. \] (14)

Combining the above equation with the facts that \( \Pi_l = (m - 1)L/\alpha \) and \( r = \rho + \dot{L}/L \), we obtain

\[ \frac{\dot{L}}{L} = \chi \left[ \left( m - 1 + \frac{1}{\alpha} \right) L - \frac{1}{\alpha} \right] - \rho. \] (15)

Equation (15) defines the evolution of the economy around the \( H \) steady state, together with the condition

\[ \dot{\alpha} = \tau (1 - \alpha). \] (16)

The \( H \) steady state is then characterized by

\[ \alpha_H^* = 1; \] (17)

\[ L_H^* = \frac{\rho + \chi}{\chi m}; \] (18)

\[ \lambda_H^* = \frac{(m - 1) \chi - \rho}{\chi m}. \] (19)

Linearizing the system, (15) and (16), we can show that the \( H \) steady state is a saddle. The non-negativity of \( \lambda_H^* \) requires that \( m > 1 + \frac{\rho}{\chi} \).

The value to a leader who is hypothetically two steps ahead in the \( H \) steady state can be computed as

\[ V_{l,2}^i = \frac{(m - 1) \rho + \chi + \tau}{\chi (\rho + \tau)}. \]

We also note that, to guarantee that indeed leaders do not want to innovate, so that they optimally choose \( \lambda_{l,2}^* = 0 \), we must have \( m < \overline{M} \), where \( \overline{M} \) is defined so that \( V_{l,2}^i = 2/\chi \). One can show that

\[ \overline{M}(\tau) = \frac{\rho + \chi}{\chi - \rho - \tau}. \] (20)

For the above to be a meaningful condition, we assume \( \chi > \rho + \tau \).

---

\(^6\) For \( m < 1 + \frac{\rho}{\chi} \), the steady state will be characterized by \( \alpha_H^* = L_H^* = 1 \) and \( \lambda_H^* = 0 \), a case that we rule out for its lack of relevance.
3.2 The M and L steady states

In the M and L steady states, both leaders and followers innovate in the contestable state. The first order conditions for $\lambda^l$ and $\lambda^f$ give $V_2^l(t) = 2\omega_t/\chi = 2V_1^l(t)$. Substituting these conditions into the value functions and after appropriate calculations we obtain the two equations:

$$
\begin{aligned}
\Pi &= (2\lambda^f - \frac{\tau}{\chi})\omega \\
\dot{\omega} &= (r + \tau - \chi \lambda^f)\omega.
\end{aligned}
$$

(21)

The M and L steady states differ in whether the maximum supply of R&D labor is exhausted or not. In the M steady state the industry-level R&D labor supply constraint is not binding. Hence, $\lambda^l_M + \lambda^f_M = \Lambda^*_M < \Lambda$ and $\omega^*_M = 1$. In contrast, in the L steady state, the industry-level skilled labor supply binds at $\Lambda$ and $\omega^*_L > 1$. In either steady state, the interest rate is $\rho$ and the second equation of (21) implies followers innovate at the intensity

$$
\lambda^f_M = \lambda^f_L = \frac{\rho + \tau}{\chi}.
$$

(22)

In the M steady state, we can solve out the production labor from the first equation in (21), together with $\Pi^*_M = (m - 1)L^*_M$ and $\omega^*_M = 1$:

$$
L^*_M = \frac{2\lambda^f_M - \tau}{m - 1} = \frac{2\rho + \tau}{\chi(m - 1)}.
$$

(23)

In the L steady state, since the skilled labor supply binds, we have $\lambda^l_L = \Lambda - \frac{\rho + \tau}{\chi}$. The evolution of the extensive margin (Eq. (12)) then implies that at the steady state,

$$
\alpha^*_L = \frac{\tau}{\tau + \chi \lambda^l_L} = \frac{\tau}{\chi \Lambda - \rho}.
$$

(24)

This, in turn, pins down production labor and profit

$$
L^*_L = 1 - \alpha^*_L \Lambda = 1 - \frac{\tau \Lambda}{\chi \Lambda - \rho};
$$

$$
\Pi^*_L = (m - 1)L^*_L.
$$

Note that for $L^*_L$ to be an interior solution to the feasibility set $[0, 1]$ we must have

$$
0 < \tau < \chi - \frac{\rho}{\Lambda}.
$$

(25)

---

7 For detailed derivations, see Appendix A.2.
With these inputs, we can solve out from (10) the value to a leader who is one step ahead,

$$V^{l*}_{1,L} = \frac{\Pi^*_L}{\rho + \chi \lambda^*_L} = (m - 1) \frac{(\chi - \tau)A - \rho}{(2\rho + \tau)(\chi A - \rho)}.$$ 

For the $L$ steady state to exist, $V^{l*}_{1,L} = \frac{\omega^*_L}{\chi} > \frac{1}{\chi}$. Define $M$ so that $V^{l*}_{1,L} = \frac{1}{\chi}$ for $m = \underline{M}$. Therefore, the existence of the $L$ steady state requires $m > \underline{M}$, which is given by

$$M(\tau) = 1 + \frac{(2\rho + \tau)(\chi A - \rho)}{\chi [(\chi - \tau)A - \rho]}.$$ (26)

The feasibility condition $0 \leq \lambda^*_L = \frac{\omega^*_L}{\chi}$ automatically guarantees that $M(\tau) \leq \overline{M}(\tau)$.

### 3.3 Discussion

The results derived in the previous section can be collected as follows.

**Proposition 1** There are two constants $\underline{M} < \overline{M}$, defined by (20) and (26), such that

i) For $m < \underline{M}$, only the $H$ steady state exists. For $m > \overline{M}$ only the $L$ steady state exists. For $m \in [\underline{M}, \overline{M}]$ the steady states $H$, $M$, $L$ all exist. Over the interval where the $M$ steady state exists, $\alpha^*_M$ is increasing in $m$.

ii) The $H$ and the $L$ steady states have the saddle-path property, while the $M$ steady state is a source. In particular, if $m > \overline{M}$, then, for any initial condition $\alpha_0$, the economy always converges to the $L$ steady state.

**Proof** See Appendix A.2.

For the range of parameters where all three steady states exist, the steady states are ranked by their respective extensive margin of innovation, $\alpha^*_H > \alpha^*_M > \alpha^*_L$, and hence our naming of the steady states. Furthermore, the following corollary to Proposition 1 makes it clear that the steady states are also ranked by their respective equilibrium growth rate.

**Corollary 1** For $m \in (\underline{M}, \overline{M})$, the steady state growth rates satisfy $g^*_H > g^*_M > g^*_L$. Moreover, the steady state growth $g^*_L$ associated with $m > \overline{M}$ is smaller than the growth rate in any of the steady states for $m \in (\underline{M}, \overline{M})$.

**Proof** See Appendix A.3.

It is worth emphasizing that the multiplicity of steady states emerges only after we relax sufficiently the assumption of instantaneous knowledge spillover in the GH setting. Equation (25) tells us that if knowledge spillovers as represented by $\tau$ are not
Innovate to lead or innovate to prevail... 881
too large (certainly not infinity), then an \( L \) steady state exists for high enough values of \( m \). In the case of knowledge spillovers that instead are at least as large as \( \tau \geq \chi - \rho \), our model features only the \( H \) type. That is, our paper nests exactly the GH model for \( \tau \geq \chi - \rho \) (and \( \bar{\Lambda} \geq 1 \)).

The mechanism behind the structure of the steady states is based on the joint effect of two simple properties of the model: the effect of leaders’ innovation on aggregate growth, and the effect of the net present value of monopoly on leaders’ incentives to innovate.

First, greater innovation by leaders is associated with lower long-run technological growth—the result outlined in Corollary 1. This is not surprising, since leaders’ innovation is motivated by endpoint strategies, whose sole goal is to discourage innovation. The positive effect on growth from a higher intensive margin of innovation carried out by leaders in contestable states is more than offset by the greater fraction of industries that, in the long run, end up in non-contestable states with zero innovation. To see this, recall that the rate of technological growth is proportional to the product of the extensive and intensive margins. Using (12) to calculate the steady state value of the extensive margin, we obtain

\[
g^* = \log(\gamma) \chi \alpha^* \Lambda^* = \log(\gamma) \frac{\tau}{\tau + \chi \lambda_f^*} \Lambda^* = \log(\gamma) \frac{\tau}{\tau + \chi \lambda_f^*} (\lambda_f^* + \lambda_l^*).
\]

A larger innovation rate by leader is then associated with lower aggregate growth, provided that \( \chi \lambda_f^* > \tau \). This latter condition, which in our case always holds in equilibrium, simply requires that the probability of a successful innovation by followers in the contestable state is greater than the exogenous probability of a spillover (or of a patent expiration) in the non-contestable state.

Second, a higher net present value of monopoly power increases leaders’ incentives to innovate. This is intuitive, since a larger leadership value triggers more effort to secure it by means of endpoint strategies. Indeed, it is straightforward to show that the incremental value \( V_l^2 - V_l^1 \) to a leader who successfully innovates is given by

\[
V_l^2 - V_l^1 = \frac{\chi \lambda_f^*}{(r + \chi \lambda_f^*)(r + \tau)} \Pi.
\]

Leaders’ endpoint strategies are incentivized when monopolist’s profits \( \Pi \) are higher, when followers’ innovation rates \( \lambda_f \) are higher, when interest rates \( r \) are lower, and when the externality intensity \( \tau \) is lower. Profits are higher, for instance, when markups \( m \) are larger, which explains the result of Proposition 1, depicted in Fig. 1. Panel (a) of the figure plots the steady state extensive margin of innovation \( \alpha^* \) against \( m \). As the figure shows, if markups \( m \) are too large, i.e. \( m > \bar{M} \), then incentives for leaders

\[\text{8} \text{ This can also be seen from the fact that as } \tau \uparrow \chi - \rho \text{ then } M(\tau) \to \infty. \text{ If } \tau > \chi - \rho, \text{ then the } L \text{ steady state still disappears but the first order conditions used to characterize it are incorrect because they were obtained under the implicit assumption of an interior solution. Consequently, all the equilibrium quantities derived above for the bounds } \bar{M} \text{ and } \bar{M} \text{ lose meaning.} \]

\[\text{9} \text{ At the steady state, (10) implies } V_l^1 = \frac{\Pi}{r + \chi \lambda_f^*} \text{ and (8)-(10) implies } V_l^2 - V_l^1 = \frac{\chi \lambda_f^* V_l^1}{r + \tau}. \text{ Combining the two, we obtain the expression for } V_l^2 - V_l^1. \]
Fig. 1 Extensive margin of innovation and growth rate in the steady states. Note: This figure illustrates the structure of the steady states of the model. In particular, it shows how the steady state extensive margin of innovation $\alpha^*_i$ and the steady state growth rate $g^*_i$ vary as we vary the markup parameter $m$. For a discussion, see Sect. 3.3

to innovate are so strong that for any initial condition the economy converges to $L$, which is the steady state with lowest long-run growth (Panel (b)). A similar reasoning explains why when knowledge spillovers are infrequent, then only the $L$ steady state exists (notice that $\bar{M}$ is increasing in $\tau$). Instead, a higher discounting $\rho$ reduces the present value of profits, and thus discourages R&D by leaders, while higher innovation
by followers, by increasing the threat to the incumbents, strengthens their incentives to play end point strategies.

For values of $m \in [M, \bar{M}]$, the model features multiple steady states. The key to understanding this result is to link, in general equilibrium, the two properties discussed above. Fix a given $m \in [M, \bar{M}]$ and begin by assuming that the economy is in a high growth steady state. Since a large fraction of the labor input is devoted to R&D, production labor and period profits are low. With low profits, incentives for leaders to innovate fall. For similar reasons, innovation intensity for followers $\lambda_f^*$ is also reduced (in an $H$ steady state the extensive margin of innovation is large, but the intensive margin is small). This further depresses leaders’ innovation incentives. In conclusion, if the economy is at a $H$ steady state, then general equilibrium effects discourage innovation by leaders, and since R&D by leaders is negatively associated with long run growth, then the high growth steady state is self-confirmed. A similar line of reasoning can be followed, for instance, to self-confirm an initial position at the $L$ steady state. In particular, in the region with multiple steady states, the $L$ equilibrium represents a “monopolistic growth trap.”

4 Global dynamics

We have established that the model contains three steady states, two of which are saddle-path stable while the other is unstable. One may now wonder whether there exist also other types of equilibria, such as cycles where leaders oscillate between innovating and not innovating. Proposition 2 establishes that no such equilibrium exist. This result indicates that the non-existence of cycles in the basic GH model does not hinge on the assumption, that we here relax, of large knowledge spillovers.

Proposition 2 Consider any equilibrium where variables are continuous at all $t$, with the possible exception of times where $\lambda^*_l$ drops to zero. Then for $t \to \infty$ the economy must converge to one of the three steady states $L$, $M$, or $H$.

Proof See Appendix A.4.

Note that Proposition 2 holds also for equilibria where the variables are allowed to change discontinuously at specific points. In this way, Proposition 2 encompasses the case of equilibrium paths that converge to the $H$ steady state but that start at time 0 with strictly positive innovation rates $\lambda^*_0 > 0$ by leaders. These equilibria can in fact feature a point of discontinuity caused by the linearity of the R&D technology, which

10 If we compare (22) with (19) we note in fact that $\lambda^*_H < \lambda^*_M = \lambda^*_L$ for $m < \bar{M}$.

11 Under the log-utility assumption here the steady state interest rate $r = \rho$ is independent of the steady state growth rate $g$. We show in Appendix E that, when the elasticity of intertemporal substitution is greater than 1, then a larger growth rate raises the equilibrium interest rate, further dampening innovation incentives for leaders, re-enforcing the self-confirming nature of the multiple steady states.

12 We speculate that this conclusion is likely to hold also for Romer (1990), who employs a modelling structure and delivers equilibrium relations that are relatively similar to GH. On the contrary, the different set of assumptions on which Aghion and Howitt (1992) is based, allows their model to feature equilibrium cycles.
induces a “bang-bang” structure for the leaders’ optimal innovation strategy with $\lambda^*_l$ jumping to zero at the time $t$ when leaders’ innovation ceases (incidentally, this causes the value function $V^l_2$ to display a kink at such points).

### 4.1 Saddle path dynamics: simulations

In this section, we provide an empirically plausible parametrization of the model and numerically simulate the saddle path convergence of the economy to either the $H$ or the $L$ steady state depending on the initial condition $\alpha_0$.

We begin by considering an initial condition $\alpha$ in the interval $(\alpha^*_L, \alpha^*_M)$ but close to $\alpha^*_M$, and we construct the saddle path equilibrium that leads the economy to converge over time to the $L$ steady state. This is obtained as follows. First, as the economy is initially close to the $M$ steady state, both leaders and followers innovate and the intensive margin constraint $\lambda$ is not binding. Therefore the system $(\alpha_t, L_t)$ evolves as follows

$$
\dot{\alpha}_t = \rho \alpha_t + \frac{2\rho + \tau}{m-1} - \chi L_t = L^*_M = \frac{2\rho + \tau}{\chi(m-1)}.
$$

(27)

As soon as $\lambda$ binds, the system switches to

$$
\dot{\alpha}_t = \tau - \alpha_t \left( \chi \lambda - \rho \right)
$$

$$
L_t = 1 - \alpha_t \lambda
$$

and the economy converges to the $L$ steady state.

Alternatively, we can construct a convergence path for the economy starting from an initial $\alpha_0 \in (\alpha^*_M, \alpha^*_H)$ that is sufficiently close to $\alpha^*_M$. The system evolves according to (27) until leaders are indifferent between innovating and not innovating, i.e. until $V^l_2(t) = 2/\chi$. After that, leaders stop innovation and the equilibrium jumps to the saddle path that converges to the $H$ steady state where only followers innovate. The system evolves according to:

$$
\dot{\alpha}_t = (1 - \alpha_t) \tau
$$

$$
\dot{L}_t/L_t = \chi \left( (m - 1) L_t - \frac{1-L_t}{\alpha_t} \right) - \rho.
$$

We simulate the model under the parameters given in Table 1. A period in the model is one year. We set the subjective discount rate to 0.02. The skilled labor supply cap, $\lambda$, is calibrated to the percentage of college graduates among adult population in the US in 2017, not all of whom need work in the R&D sector. The rate of patent expiration, $\tau$, is taken to be 0.05, which is the inverse of the term of patents (20 years). We set $m$ to be in the interval $[M, \bar{M}]$, so all three steady states exist. The step size $\gamma$ of a successful innovation is chosen to ensure reasonable growth rates of the economy.

The saddle equilibrium path that converges to the $L$ steady state is illustrated in Panel (a) of Fig. 2. The initial condition of the extensive margin of innovation, $\alpha_0$,
Innovate to lead or innovate to prevail...

Fig. 2
Equilibrium saddle paths. Note: This figure shows the simulated equilibrium paths to the $L$ and the $H$ steady state. In particular, we plot the evolution of the extensive margin of innovation $\alpha_t$, the amount of production labor $L_t$, and the steady state growth rate $g_t$ on their saddle paths to the steady state. The model is simulated under the parametrization given by Table 1. For a discussion, see Sect. 4.1
Table 1  Parameter values

| Parameter | Value | Justification |
|-----------|-------|---------------|
| $\rho$    | 0.02  | Convention    |
| $\bar{\lambda}$ | 0.34  | Pct. of college graduates among adult population |
| $\tau$    | 0.05  | Term of patents, 20 years |
| $\chi$    | 0.24  | Ensure the existence of path to the $H$ |
| $m$       | 1.53  | The average of $\bar{M}$ and $\bar{M}$ |
| $\gamma$  | 1.60  | Ensure $\gamma \geq m$ and reasonable consumption growth rate |

This table reports the parameter values we use in the simulation of the baseline model and their justifications. For a discussion, see Sect. 4.1

is chosen to be just below $\alpha^*_M$. Initially, around the $M$ steady state, both leaders and followers innovate. While the extensive margin of innovation decreases over time, the intensive margin of innovation increases due to increasing innovation rates by leaders. As long as the skilled labor supply is not binding, the evolution of the two counteract each other perfectly, so that the rate of technological growth pinned down by the aggregate level of innovation, $g_t$, is constant and so is the size of the production sector, $L_t$. As soon as the skilled labor supply binds in contestable industries, the decline in the extensive margin takes over, the aggregate level of innovation declines and the production sector expands. Admittedly, the kink in $g_t$ is driven by the fixed supply of skilled labor together with the linearity of the R&D technology at the firm level. In Appendix F, we show that if we relax the assumption of the linear R&D cost structure, then $g_t$ evolves smoothly as the economy converges to the $L$ steady state.

Panel (b) of the same figure shows the saddle equilibrium path that converges to the $H$ steady state. Starting from an initial extensive margin of innovation just above the $M$ steady state, the extensive margin of innovation increases, the intensive margin innovation decreases due to leaders innovating less, while the aggregate innovation stays constant, until the moment when leaders no longer find it profitable to innovate. At that point, the equilibrium jumps to the saddle path that converges to the $H$ steady state. The extensive margin $\alpha_t$ moves continuously, though its rate of change has a kink at that point, while the production labor $L_t$ has a discontinuous jump. Once this point of discontinuity is crossed, the extensive margin keeps increasing until all industries become contestable and the economy approaches the $H$ steady state.

5 Policy and welfare implications

In this section, we explore the policy and welfare implications of the model.

5.1 Policy implications

Should Medicare be allowed to bargain for better deals with drugs providers, effectively reducing the markup for pharmaceutical companies? What is the effect of longer patents’ duration? These are all common policy questions that, in our model, involve
setting the parameters \( m \) and \( \tau \). From this perspective, the main implication of Proposition 1 is that relatively small changes in \( m \) and \( \tau \) can trigger dynamic adjustments that lead to large changes in the growth rate in the final steady state. In such cases, simply reversing the change in the parameters \( m \) and \( \tau \) may not be enough to dynamically move the economy back toward its original starting point. Apparently well-meaning policies with adverse long-run implications can be very costly to reverse.

**Corollary 2** Suppose that, given parameters \( m \) and \( \tau \), the economy is at an initial steady state. Consider an unexpected and permanent change \( \Delta m \) and \( \Delta \tau \) in the parameters. Then,

i) If \( H \) is the initial steady state, then the economy transitions toward an \( L \) steady state if and only if \( \Delta m > M(\tau + \Delta \tau) - m \),

ii) If \( L \) is the initial steady state, then the economy transitions toward an \( H \) steady state if and only if \( \Delta m < M(\tau + \Delta \tau) - m \),

where \( M(\cdot) \) and \( M(\cdot) \) are given respectively by (20) and (26).

Corollary 2 can be established by looking at Panel (a) of Fig. 1. Suppose that the economy is initially at an \( H \) steady state, which implies that \( m < M(\tau) \), and assume for simplicity that \( m \) changes by an amount \( \Delta m \) while \( \tau \) remains unchanged. If \( \Delta m < M(\tau) - m \), then the economy is still at an \( H \) steady state after the policy change. But growth immediately jumps to a different steady state level along the \( g^*_H \) curve marked by squares in Panel (b) of Fig. 1. If instead \( \Delta m > M(\tau) - m > 0 \), then the \( H \) steady state disappears and the economy converges dynamically to the \( L \) steady state. In a similar fashion, Fig. 1 can be used to establish what happens when the initial steady state is \( L \) or when \( \tau \) is also changed.

An interesting feature of the dynamic path associated with a policy-induced steady state change from \( H \) to \( L \) is that the economy’s growth rate follows a boom-bust trajectory during transition: as the degree of protection of monopolistic profits is strengthened (e.g. \( \Delta m > 0 \) and/or \( \Delta \tau < 0 \)), growth initially increases before finally collapsing. A simple way to see this analytically is to consider the case where \( m \) and \( \tau \) are such that \( m = M(\tau) \) and the economy is initially at the \( H \) steady state. By Corollary 2, any policy change \( \Delta m > 0 \) and/or \( \Delta \tau < 0 \) would set the economy on a dynamic path to the \( L \) steady state. On such a path, and until the moment when \( \bar{\lambda} \) binds, we have \( L_t = L^*_M \), with \( L^*_M \) given by (23) for new policy parameters \( m' = m + \Delta m \) and \( \tau' = \tau + \Delta \tau \). By calculating \( L^*_H \) from (18) for the initial value \( m = M(\tau) \), we obtain \( \Delta L^*_H \) as the jump in production labor at the time of the policy change. Correspondingly, the initial jump \( \Delta g^*_0 \) in the economy’s growth rate is

\[
\Delta g^*_0 = -\log(\gamma)\chi \Delta L^*_0 = \log(\gamma) \frac{1}{M(\tau)} \Delta m + \frac{(2\rho + \chi)M(\tau)^2}{(\rho + \chi)^2} (\Delta m - \Delta \tau) > 0.
\]

The equation above shows that, even though the policy change causes the growth rate of the economy to fall asymptotically from \( g^*_H \) to \( g^*_L \), the initial growth rate jump

13 See Eq. (31) in Appendix A.2. Since \( \dot{\lambda}_f = 0 \) along the convergence path to the \( L \) steady state, and since \( \dot{\omega}_t = 0 \) until \( \bar{\lambda} \) binds, then also \( \dot{L}_t = 0 \) and correspondingly \( L_t = L^*_M \) until \( \bar{\lambda} \) binds.
\( \Delta g_0^* \) is strictly positive. Panel (a) of Fig. 3 illustrates an example, where the economy is initially at the \( H \) steady state of the model calibrated according to Table 1. At the instant \( t = 2 \), we increase the markup by 1 percentage point which raises \( m \) above \( \overline{M} \). The growth rate jumps up from the initial 3.28\% to 3.40\% and, as the constraint \( \Lambda_t \leq \overline{\Lambda} \) does not bind for several periods, the growth rate stays at 3.40\%, before eventually declining towards the \( L \) steady state at 3.11\%. A policy maker who, based on the implications of \( \text{GH} \), decided to stimulate growth by awarding more monopoly power and increasing the markup ceiling would feel vindicated in the short run. Growth initially jumps up. However, after some time, the policy maker would realize that the economy is actually heading toward a steady state with lower long-run growth.

The dynamic adjustment of the growth rate then raises naturally the following question: can the policy maker reverse its course by undoing the initial policy? Panel (b) in Fig. 3 illustrates such an attempt. A time 50, when the economy is clearly heading the wrong direction, the policy maker could (unexpectedly) undo the policy by restoring the markup ceiling to the original level. But the simulation shows that it would have no effect on long-run growth and very small consequences for short-run dynamics. In order to set the economy again on course to a \( H \) steady state and thus alleviate the negative long-run consequences of the initial policy change, the policy maker would have to reduce the markup ceiling by more than it was raised in the first place. Panel (c) shows that if at time 50, \( m \) is decreased by 2 percentage points, growth initially drops but then the economy reverses its course and heads toward a \( H \) steady state with a growth rate of 3.22\%: this is still lower than that in the original \( H \) steady state, as \( m \) is now 1 percentage point lower.

More generally, Corollary 2 helps us to understand the “reversibility” of a policy change. When a dynamic switch from an \( H \) to a \( L \) steady state occurs, the effectiveness of a policy reversal in re-establishing a path to an \( H \) steady state depends both on the original value of \( m \) and on the value reached by \( \alpha_t^* \) at the time of the reversal. If the policy reversal is only implemented when \( \alpha_t^* \) is sufficiently close to the new \( L \) steady state, then point ii) of Corollary 2 implies that the reversal would only be effective if the reduction \( \Delta m \) is large enough that \( m \) falls at or below \( \overline{M} \). This means that, if the original policy change increased \( m \) from a level in the interval \([ M, \overline{M} ]\) to above \( \overline{M} \), then the size of the policy reversal needs to be bigger (in absolute value) than the initial policy change. Moreover, even if effective, such reversal will result in a reduction in long-run growth compared to the initial \( H \) steady state because of the over-compensating policy reversal. On the other hand, if the policy reversal is implemented shortly after the original policy change, so that the extensive margin of innovation \( \alpha_t^* \) still in the region of attraction of the original \( H \) steady state, restoring the original \( m \) can in principle fully reverse the course of the equilibrium path.\(^{14}\)

Similar to an increase in \( m \), we can perform policy experiment that reduces the length of patent protection, \( \tau \). For example, imagine we are initially at an \( H \) steady state with \( m \) close to but lower than \( \overline{M}(\tau) \). Increasing the length of patent (reducing \( \tau \)) results in \( m > \overline{M}(\tau - \Delta \tau) \), which propels the economy to switch to the \( L \) steady state.

\(^{14}\) Of course, if the initial \( m \) is below \( \overline{M} \) so that only the \( H \) steady state exists, then reversing \( m \) to the initial level any time along the dynamic path to \( L \) will restore the initial steady state.
Fig. 3  Markup policy and its reversals. Note: This figure illustrates how an economy parameterized according to Table 1 evolves when $m$ is increased by 1 percentage point in period 2 (a); when $m$ is further restored to the original value in period 50 (b); and when $m$ is instead restored to the original value minus 1 percentage point in period 50 (c). For a discussion, see Sect. 5.1.
A similar discussion of policy reversal as the one in Fig. 3 applies in this case and, for brevity, we leave the corresponding simulation results to Appendix C.

5.2 Pareto optimality

Equilibria featuring higher innovation may or may not be desirable from a welfare perspective. The optimal innovation rate depends in fact on the relation between the social benefit of a successful innovation, represented by $\gamma$, the cost of achieving a successful innovation, proxied by $1/\chi$, and the rate of time preferences, $\rho$.

The optimal innovation rate can be calculated from a planner’s problem. The social planner splits the fixed supply of labor between production and R&D activities across industries to maximize the lifetime utility of the representative consumer. The problem is

$$\max_{l_{it}, \Lambda_{it}} \int_0^\infty e^{-\rho t} \log C_t dt$$

subject to:

- $\log C_t = \int_0^1 \log (q_{it} d_{it}) di$
- $d_{it} = l_{it}$
- $q_{it} = \gamma^s_{it}$
- $\dot{s}_{it} = \chi \Lambda_{it}$
- $\int_0^1 (l_{it} + \Lambda_{it}) di \leq 1$
- $\Lambda_{it} \leq \bar{\Lambda}$.

where as usual $\Lambda_{it}$ represents the overall innovation effort in industry $i$ and the last two constraints are the labor resource constraint and the fixed skill constraint.

**Proposition 3** The solution to the planner’s problem is a unique steady state with consumption growth rate $g^{SP} = \chi \log(\gamma) - \rho$. Furthermore:

1. The decentralised H steady state features a suboptimally low growth rate $g^*_H < g^{SP}$ if and only if

$$m < \log(\gamma) \left( \frac{\chi}{\rho} + 1 \right)$$

2. The decentralised L steady state features a suboptimally high growth rate $g^*_L > g^{SP}$ if and only if

$$\tau > 1 - \frac{\rho}{\chi \Lambda} \left( \frac{\chi \bar{\Lambda} - \rho}{\chi \log(\gamma) + 1} \right)$$

**Proof** See Appendix D. \qed
In the $H$ steady state of the decentralised economy, only followers innovate and the steady state behaves similarly to the decentralised steady state in GH where in addition $m = \gamma$. Under this condition, Proposition 3 implies that if $\gamma < \log(\gamma) \left( \frac{\rho}{\rho + 1} \right)$ then the GH version of the $H$ steady state features underinvestment in innovation. On the contrary, when $\gamma > \log(\gamma) \left( \frac{\rho}{\rho + 1} \right)$ the equilibrium features overinvestment. By tracing out the shape of these inequalities as a function of $\gamma$, one can see that, depending on the other model parameters, there can exists a range of values $[\gamma_1, \gamma_2]$ where the GH equilibrium features underinvestment for $\gamma$ in that range, and overinvestment for $\gamma$ outside of that range. Intuitively, if the positive knowledge externality from innovation (which increases with the logarithm of $\gamma$) outweighs the negative externality from business stealing effects (which increases linearly with the profit markup $m = \gamma$), then there is too little growth in the decentralised economy.

Turning again to the more general formulation of our model, the leader’s markup can be taken to be a regulated value $m$ strictly lower than $\gamma$. Based on the discussion above, this means that for a given $m < \gamma$, the negative welfare consequence of the business stealing effects are smaller than in the GH case with $m = \gamma$. Therefore the interval of values for $\gamma$ where the $H$ equilibrium features underinvestment shrinks relative to the GH case. 15

In the event that the $H$ equilibrium features overinvestment, then Proposition 3 states that the $L$ steady state is characterized by overinvestment if and only if $\tau$ is sufficiently large. This is a straightforward consequence of the fact that the growth rate $g_L^*$ in the $L$ steady state is strictly increasing in $\tau$, i.e. $g_L^*$ is higher when the patent duration is shorter (see the proof of Corollary 1 in Appendix A.3). On the other hand, when $\tau$ approaches zero then $g_L^*$ also approaches zero, which is surely lower than the social optimum.

6 Conclusion

Traditional Schumpeterian models such as Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1992) deliver the result that higher monopoly power—higher markup or longer patent protection—unambiguously lead to higher aggregate growth. In this paper, we show that this conclusion rests crucially on the assumption that knowledge spillovers to followers are large. In many practical cases where this assumption fails, the cost for followers to leapfrog the industry’s leader increases in the leader’s technological advantage. We show that, when this alternative assumption is adopted, results of the traditional Schumpeterian model can change dramatically.

First of all, instead of being characterized by just one steady state where only followers innovate (an $H$ steady state), the economy now features another dynamically stable steady state where also leaders innovate (the $L$ steady state). The $L$ steady state has lower growth than the traditional $H$ steady state. It is characterized by high but infrequent innovation effort by industry leaders, whose “endpoint strategy” is to acquire new patents in order to distance themselves from the followers, thus increasing

---

15 Yet, in all the numerical example and policy experiments in the previous sections, the $H$ steady state features underinvestment in innovation relative to the social optimum.
the followers’ innovation costs and pushing them out of the innovation race. Second, we find that when leaders are granted large monopolistic rents or long-lasting patent protection, the $H$ steady state disappears and only the $L$ steady state remains.

Because of its structure of multiple steady states, the model has rich policy implications. A small increase of monopoly power, either in the form of a higher markup ceiling or longer patent protection, can cause the economy to lose its $H$ steady state and transition toward the $L$ steady state. Allowing leaders to take advantage of excessively high markups and long patent protection is harmful to long-run growth, as these conditions provide leaders with incentives to enact strategies aimed at stifling firms’ entry into their industry. The equilibrium transition path to the $L$ steady state features a non-monotone growth trajectory, with the growth rate rising in the short run followed by a long-run decline. Moreover, should the policy maker attempt to reverse the policy change over the course of decline, he may find that simply restoring the original level of monopoly power is not sufficient to restore the $H$ steady state. A larger policy correction may be needed.

Our theoretical findings indicate that some standard results of the Schumpeterian literature of endogenous growth are not robust to the relaxation of the assumption of large knowledge spillovers. Although we have mainly interpreted the extent of knowledge spillovers as governed by the patent policy in this paper, many other factors matter too. For example, firms can sign costly long-term contracts with key R&D personnel to limit knowledge spillovers (Gersbach and Schmutzler 2003) or hire lawyers to protect their monopolistic rents and raise legal barriers to entrants (Dinopoulos and Syropoulos 2007). Our results provide a general framework to interpret recent empirical trends of increasing markups, reduced investment, and lower business dynamism.

Acknowledgements We gratefully acknowledge helpful comments from an anonymous referee and from Phillipe Aghion, Michele Boldrin, Russell Cooper, Mario Crucini, Domenico Ferraro, Gino Gancia, Berthold Herrendorf, Yang Lu, Rodolfo Manuelli, Ramon Marimon, Stephen Parente, Ed Prescott, and Michael Song. We also thank seminar and conference participants at the Arizona State University, Washington University in St. Louis, LSE, PSE, Queen Mary University of London, University of Manchester, CAF, Chinese University of HK, Birkbeck Centre for Applied Macro Workshop 2019, CFM London Macro Workshop 2019, Southwestern University of Finance and Economics, RES Meetings 2019, HKUST-Jinan Conference on Macroeconomics 2019, SAET Conference 2018, Midwest Macro Meetings 2018. The views expressed here are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Declarations

Statements and Declarations No funding was received for conducting this study. The authors have no relevant financial or non-financial interests to disclose. The authors do not analyse or generate any datasets, because the work proceeds within a theoretical and mathematical approach.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Appendix

A Technical details of the baseline model

A. 1 The supply function of skilled labor

To obtain an equilibrium when the supply function for specialized labor is perfectly elastic up to \( \Lambda_1 \), and perfectly inelastic afterwards, we proceed in two steps. First, we postulate the existence of an exogenous supply function for the specialized labor given by

\[
\omega = 1 + \theta \psi(\Lambda)
\]

(28)

where \( \theta > 0 \) and \( \psi(\cdot) \) is a \( C^1 \) function such that

\[
\begin{align*}
\psi(\Lambda) &= 0, \quad \text{for } \Lambda \leq \Lambda_1 \\
\psi', \psi'' &> 0, \quad \text{for } \Lambda > \Lambda_1 \\
\psi(\Lambda) &\to +\infty, \text{ as } \Lambda \to +\infty.
\end{align*}
\]

(29)

Second, we take the limit of the resulting equilibrium as \( \theta \to +\infty \). For the purpose of this paper, we take \( \Lambda_1 \) to be an arbitrarily large constant.

A. 2 The Proof of Proposition 1

The \( H \) steady state and its existence are established in the main text of the paper. Here we focus on the \( M \) and \( L \) steady states.

From the first-order conditions of the innovating leaders and followers, we have

\[
2\omega_1 / \chi = 2V_1^l(t) = V_2^l(t).
\]

Therefore,

\[
2\dot{V}_1^l = 2\dot{\omega}_1 / \chi = \dot{V}_2^l. \tag{30}
\]

Substituting these conditions for the value functions and their derivatives into the equations defining \( V_1^l \) and \( V_2^l \), we obtain two equations. The first is

\[
\Pi_1 = (2\lambda_1^f - \tau / \chi)\omega_1,
\]

and the second equation is

\[
\dot{\omega}_1 = (r_1 + \tau - \chi \lambda_1^f)\omega_1.
\]

They comprise the system of equations of (21) in the paper.

From the first equation of (21), we have

\[
\frac{\dot{\Pi}}{\Pi} = \frac{\dot{L}}{L} = \frac{\dot{Y}}{Y} = \frac{\dot{E}}{E} = r - \rho = \frac{2\chi \dot{\lambda}_1^f}{2\chi \lambda_1^f - \tau} + \frac{\dot{\omega}}{\omega} = \frac{2\chi \dot{\lambda}_1^f}{2\chi \lambda_1^f - \tau} + r + \tau - \chi \lambda_1^f
\]

\Rightarrow \quad 2\dot{\lambda}_1^f = (2\lambda_1^f - \tau / \chi) \left( \lambda_1^f - \frac{\rho + \tau}{\chi} \right) \chi. \quad \tag{31}

From the expression for \( \Pi, 2\lambda^f - \tau/\chi > 0 \). This means that for (31) not to explode we must have \( \lambda^{f^*} = \frac{\rho + \tau}{\chi} \) at both an \( M \) and \( L \) steady state and along the saddle path to the \( L \) steady state. Since \( \Pi^* = (m - 1)(1 - \alpha^* \Lambda^*) \),

\[
(m - 1)(1 - \alpha^* \Lambda^*) = \frac{2\rho + \tau}{\chi} \omega(\Lambda^*)
\]

\[
\Rightarrow \alpha^* = \frac{1}{\Lambda^*} \left( 1 - \frac{2\rho + \tau}{\chi(m - 1)} \omega(\Lambda^*) \right) \equiv v_1(\Lambda^*).
\]

From \( \dot{\alpha} = (1 - \alpha) \tau - \alpha \chi \lambda^f \), we have

\[
0 = (1 - \alpha^*) \tau - \alpha^* \chi \lambda^{f^*} = (1 - \alpha^*) \tau - \alpha^* \chi (\Lambda^* - \lambda^{f^*})
\]

\[
= (1 - \alpha^*) \tau - \alpha^* \chi (\Lambda^* - \frac{\rho + \tau}{\chi})
\]

\[
\Rightarrow \alpha^* = \frac{\tau}{\Lambda^* \chi - \rho} \equiv v_2(\Lambda^*).
\]

The system of equations

\[
\begin{cases}
\alpha^* = \frac{1}{\Lambda^*} \left( 1 - \frac{2\rho + \tau}{\chi(m - 1)} \omega(\Lambda^*) \right) \equiv v_1(\Lambda^*) \\
\alpha^* = \frac{\tau}{\Lambda^* \chi - \rho} \equiv v_2(\Lambda^*)
\end{cases}
\]

when having two meaningful solutions, define the \( M \) and \( L \) steady states.

In the limit economy, let \( \theta \to +\infty \). Then,

\[
v_1(\Lambda^*) = \begin{cases}
\frac{1}{\Lambda^*} \left( 1 - \frac{2\rho + \tau}{\chi(m - 1)} \right) \text{ if } \Lambda^* < \Lambda \\
-\infty \text{ if otherwise}
\end{cases}
\]

In Fig. 4, we plot \( v_1 \) (for both a finite \( \theta \) and for the limit when \( \theta \to \infty \)) and \( v_2 \). One can show that for finite \( \theta \), \( v_1 \) and \( v_2 \) have at most two crossings, because \( \frac{v_1'}{v_2'} \) increases in \( \Lambda \).\(^{16}\) As \( \theta \to \infty \), the lower crossing occurring at the binding skilled labor constraint, defining the \( L \) steady state, \( \Lambda^*_L = \Lambda \) and the higher crossing defines the \( M \) steady state, where \( \omega^*_M = 1 \).

Let \( \theta \) go to infinity. Varying \( m \) shifts \( v_1(\cdot) \) up and down. Let \( M \) be the \( m \) such that \( v_1(\cdot) \) and \( v_2(\cdot) \) have only one intersection at \( \Lambda \). This implies that if \( m \) is lower than \( M \), then the \( M \) and \( L \) steady states disappear. Let \( \overline{M} \) be the \( m \) such that there are two

\[\text{We have}
\frac{v_1'}{v_2'} = \frac{2\rho + \tau}{\chi(m - 1)} \theta \psi' \Lambda + \frac{v_1 \Lambda}{\tau} \left( \chi - \frac{\rho}{\Lambda} \right)^2.
\]

It can be shown that the first term’s derivative with respect to \( \Lambda \) is \( \frac{2\rho + \tau}{\chi(m - 1)} \Lambda \psi'' > 0 \). The second term is clearly increasing in \( \Lambda \). Therefore, overall the ratio \( \frac{v_1'}{v_2'} \) is increasing in \( \Lambda \).
intersections of $v_1(\cdot)$ and $v_2(\cdot)$, with the higher one corresponding to $\alpha^*_H = 1$ and the lower one corresponding to $\Lambda$. This implies that if $m = \overline{M}$, then $\alpha^*_H = \alpha^*_M = 1$. If $m > \overline{M}$, then the H and M steady states disappear. We can show that

\[
\overline{M} = \left\{ \begin{array}{ll}
\frac{x + \rho}{\chi - \tau - \rho} & \text{if } \chi - \tau - \rho > 0 \\
+\infty & \text{if otherwise.}
\end{array} \right.
\]

In sum, when $m < \underline{M}$, there is only one H steady state.

When $\underline{M} \leq m \leq \overline{M}$, there are three steady states, $H$, $M$, and $L$.

\[
\begin{align*}
\alpha^*_H &= 1; \\
\alpha^*_M &\to \frac{x - \tau - 2\rho + \tau}{m - 1}; \\
\alpha^*_L &\to \frac{\tau}{\chi \Lambda - \rho}.
\end{align*}
\]

In the $M$ steady state, $\Lambda^*_M < \overline{\Lambda}$ and $\omega^*_M = 1$. In the $L$ steady state, $\Lambda^*_L = \overline{\Lambda}$ and $\omega^*_L > 1$.

When $m > \overline{M}$, there is only one $L$ steady state.

The stability properties of the $M$ and the $L$ steady state are easily established. In the jargon of economy theory, a steady state is locally stable if, given an initial condition
(in our case, an initial value for $\alpha(0)$ in the neighborhood of the steady state, there exists an equilibrium path converging to the steady state as $t \to \infty$. We know that in a neighborhood of either the $M$ or the $L$ steady state, the differential equation (31) must hold. In order to have $\lambda^f_i$ converge to its steady state value $\lambda^f_* = \frac{\rho + \tau}{\chi}$, we must necessarily have $\lambda^f_i = \frac{\rho + \tau}{\chi}$ for all $t$. Hence, starting from any $\alpha_0 \in (\alpha^*_L, \alpha^*_M)$, there exists a unique equilibrium pair $(\alpha^*_L, \alpha^*_M)$, which travels southeast along the curve $v_1$ and converges to the $L$ steady state as $t \to \infty$. Similarly, pick any $\alpha_0 < \alpha^*_L$. We have $\alpha_0 > \alpha_L^*$ along the trajectory, which implies that there exists a unique equilibrium pair $(\alpha_0, \lambda_0)$ which travels northwest along $v_1$ and converges to the $L$ steady state. Finally, for any $\alpha_0 > \alpha_L^*$ we have $v_1 < v_2$. Therefore, any path starting and lying on $v_1$ is characterized by $\alpha_0 > \alpha^*_L$ for all $t$, which shows that there is no initial condition $\alpha_0$ in the neighborhood of $\alpha^*_M$ for which we can find an equilibrium path converging to the $M$ steady state. We then say that $M$ is a source and $L$ is a saddle.

Clearly, if $m > \overline{M}$ then the only steady state is $L$, and for any initial condition $\alpha_0 \in (0, 1]$ the only equilibrium is the one associated with the unique path converging to the $L$ steady state.

A. 3 The Proof of Corollary 1

Firstly, note that aggregate R&D labor in the $M$ and $L$ can be expressed by

$$\alpha^*_i \lambda^*_i = \frac{\tau \Lambda^*_i}{\tau + \chi (\Lambda^*_i - \lambda^f_i)} \mbox{, for } i = M, L.$$ 

Since $\lambda^f_M = \lambda^f_L$ and $\Lambda^*_M < \Lambda^*_L = \overline{\Lambda}$, we conclude that $\alpha^*_M \Lambda^*_M < \alpha^*_L \Lambda^*_L$ under the assumption that $\tau < \chi - \rho < \chi$. Since $g^*_i = \log(\gamma) \chi \alpha^*_i \Lambda^*_i$, we have $g^*_M > g^*_L$.

---

17 Recall that necessarily $2\lambda^f_i - \tau / \chi > 0$. If $\lambda^f_i$ increases to the steady state value, then $\lambda^f_i > \frac{\rho + \tau}{\chi}$ and $\lambda^f_i$ will increase without bound. If $\lambda^f_i$ decreases to the steady state value, then $\lambda^f_i < \frac{\rho + \tau}{\chi}$ and $\lambda^f_i$ will decrease to zero. Either is a contradiction.
Secondly, comparing the aggregate production labor in the $H$ and $M$ steady states, (18) and (23), we find $L^*_H < L^*_M$ if and only if $m < \bar{M}$. Since $g_i^* = \log(\gamma)\chi(1 - L_i^*)$, we have $g_H^* > g_M^*$ when both exist. Substituting (18) in $g_H^*$, it is easy to see that $g_H^*$ is increasing in $m$ and independent of $\tau$.

Finally, we can easily solve out the growth rate in the $L$ steady state: $g_L^* = \frac{\log(\gamma)\chi}{\Lambda_1\chi/\Lambda_1 - \rho}$, which is independent of $m$ (and is increasing in $\tau$). Therefore, $g_L^*$ is smaller than any growth rates in the $M$ and $H$ steady states for any $m \in (\bar{M}, \bar{M})$. This proves Corollary 1.

A. 4 The Proof of Proposition 2

Define “Region I” the system of differential equations (15)–(16) where only followers innovate and “Region II” the system (30)–(31) where both leaders and followers innovate.

Assume that the initial conditions of the system do not coincide with either the steady states or the saddle paths of the two Regions. Then, starting from such initial conditions, a candidate equilibrium must necessary switch Region at least once, possibly featuring equilibrium cycles where the economy switches indefinitely between Regions (it is straightforward to show that within-Region cycles do not exist).

Indicate with $T_1$ a time when the economy switches Region and consider a candidate equilibrium characterized by an infinite sequence $T_1, T_2, \ldots$ of switches between Regions. For brevity, in the remainder of the proof we assume that $\alpha(0) > \alpha_L^*$ and $\lambda^f(T) < \bar{\lambda}$, implying $\omega(T) = 1$ at any $T$. Also, with a slight abuse of notation, we will indicate with $x(T)$ or $\dot{x}(T)$, respectively, the limit from the right of the function $x(T)$ or of its time derivative.

Without loss of generality we assume that at $T = T_1$ the economy switches from Region I to Region II. Because of the continuity assumption, we must have $\lambda^l(T) = 0$. There are now two possibilities that need to be considered separately.

The first is that $\lambda^f(T_1) \geq \lambda_M^f = \frac{\beta + \tau}{\chi}$. In this case (31) implies that $\dot{\lambda}^f(T_1) \geq 0$. Also, since $\Pi = (m - 1)L$, equation (A-3) implies $\dot{L}(T_1) \geq 0$. Moreover, since $\lambda^l(T_1) = 0$, then $\dot{\alpha}(T_1) > 0$. Given that $\lambda^l_t = \frac{1 - L_t}{\alpha_t} - \lambda^f_t$ for any $t$, it follows that $\dot{\lambda}^l(T_1) < 0$. But this is not possible, since it would imply that $\lambda^l_t < 0$ at some time $t > T_1$.

For the rest of the proof we will then focus on the second possibility, that is $\lambda^f(T_1) < \lambda_M^f$. Since $\Pi_t = (m - 1)L_t$, then using equation (30) we have

$$\frac{\dot{L}}{L} = \frac{\dot{\Pi}}{\Pi} = \frac{2\lambda^f}{2\lambda^f - \frac{\beta + \tau}{\chi}}.$$  

Substituting (31) gives

$$\frac{\dot{L}}{L} = \chi\lambda^f - \rho - \tau.$$  

\[\square\]
Recalling that $\lambda^f(T_1) < \lambda^f_M$, we conclude that $\dot{L} < 0$ and $\dot{\alpha}_1 < 0$ while the system is in Region II after the switch at $T_1$.

Then, at the time $T_2 > T_1$ when the economy switches back to Region I, we have $\lambda^f(T_2) < \lambda^f(T_1), L(T_2) < L(T_1), \dot{\lambda}^l(T_2) = 0$ and thus $\alpha(T_2) = \frac{1 - L(T_2)}{\lambda^f(T_2)} > \alpha(T_1)$. Since $\alpha$ strictly increases while in Region I, we conclude that as the economy switches back and forth between the two regions it generates a strictly increasing sequence $\{\alpha(T_n)\}$. This sequence can only converge to $\alpha_H^*$ given that $\alpha(0) > \alpha_H^*$ and that in Region I $\alpha > 0$ as long as $\alpha < \alpha_H^*$.

We therefore conclude that there is no equilibrium that features a cycle where the economy switches between Region I and Region II. Any equilibrium, if it exists, that at some time $T_1$ switches from Region I to Region II at some $T_1$ must converge in the limit to the $H$ steady state.

**B. Dynamic race with endogenous steps**

Assume that the innovation technologies used by followers are linear and that innovation costs for followers are increasing in the follower’s lag from the leader. We will show that, under these assumptions, the model with an exogenous maximum distance of two steps is in fact the equilibrium result of a model where the maximum distance is endogenous.

To prove the claim, we need to consider only the problem of the follower, taking as given the value functions $V^s_l$ of a leader $s$ steps ahead, for $s = 1, 2, \ldots$. Also, to simplify the exposition, we can focus only on steady states. The value $V^s_f$ of a follower $s \geq 2$ steps behind and who innovates to leapfrog the leader is given by the solution to

$$rV^s_f = \max_{\lambda^s_f \geq 0} -\lambda^s_f + (\chi_s\lambda^s_f + \tau_s)(W - V^s_f) + \lambda^l_s(V^s_{s+1} - V^s_f).$$

(34)

where $W \equiv V^1_l$. Innovation costs $1/\chi_s$ are assumed to be strictly increasing in the lag $s$, while the spillover intensity $\tau_s$ is assumed to be a decreasing sequence for $s \geq 2$. We employ the normalizations $\chi_1 = 1$ and $\tau_1 = 0$.

For brevity, we can appeal to an intuitive argument that, since innovation costs are increasing in the follower’s lag, and spillover’s intensities are decreasing, then $V^s_f \geq V^s_{s+1} \geq 0$, i.e. the follower is (weakly) better off when the distance from the leader is smaller. Consider now the case $s = 1$, where the follower’s value function

---

18 Note that we make no assumption about the form of the innovation costs of the leader.

19 The normalization $\tau_1 = 0$ is adopted simply to replicate one basic tenet of the Schumpeterian literature: innovation by followers is necessary to displace the leader in the product market. In this sense, we can say that technological knowledge spillovers from leaders to followers are full at $s = 1$ (the R&D technology is the same for both firms), but this has no impact on the leader’s dominance in the product markets, where the leader continues to earn monopolistic profits $\Pi_1$ until he is actively leapfrogged by followers).

20 The result that the value to the follower decreases with the lag is standard in models of races (see for instance Hörner 2004). For the sake of our demonstration, we can make the (incorrect) assumption that $V^s_f \geq V^s_{s+1} \geq 0$. Then, optimality of (34) for $s = 2$ requires that $\lambda^s_2 = 0$. Moreover, since $V^2_f \geq 0$ and
solves the usual problem,

\[
 r V^f_1 = \max_{\lambda_1^f \geq 0} -\lambda_1^f + \lambda_1^f (W - V^f_1) + \lambda_1^f (V^f_2 - V^f_1). \tag{35}
\]

Because of the linearity of the R&D technology and since the optimal value of \(\lambda_1^f\) must be finite in equilibrium, then, regardless of whether the condition \(\lambda_1^f \geq 0\) is binding in the maximization of (35), we have

\[
 V^f_1 = \frac{\lambda_1^f}{r + \lambda_1^f} V^f_2.
\]

Since at a steady state \(r > 0\) and \(\lambda_1^f \geq 0\), the equation above implies that \(V^f_1 \leq V^f_2\).

But since we also have posited that \(V^f_1 \geq V^f_2 \geq 0\) then we must have

\[
 V^{f*}_1 = V^{f*}_2 = 0.
\]

Iterating on the conditions \(V^f_s \geq V^f_{s+1} \geq 0\) for \(s = 3, 4, \ldots\) yields

\[
 V^{f*}_s = 0, \forall s \geq 1.
\]

In any equilibrium where \(\lambda^{f*}_s > 0\), the linearity of the R&D technology and the fact that \(V^{f*}_s = 0\) implies the first order condition \(W = \frac{1}{\chi_s}\). Therefore, if this condition holds for \(s = 1\) so that \(\lambda^{f*}_1 > 0\), it cannot hold for any other \(s > 1\), since at those states \(\chi_s < \chi_1\). In conclusion, if in equilibrium followers have a strictly positive innovation rate when they are one step behind the leader, they do not innovate when they are more than one step behind. This concludes the demonstration that our assumption in the paper that followers do not innovate at state \(s = 2\) is without loss of generality.

The result presented above hinges crucially on the linearity of the production technology, which guarantees that followers earn zero profits in expected terms (\(V^{f*}_s = 0\)). In turn, the zero profit condition rests on the free-entry assumption for the competitive fringe of followers. This assumption is standard in the traditional literature of Schumpeterian models of leapfrogging innovation to which our paper contributes directly. Indeed, we can show that, by maintaining the assumption of a linear R&D technology, the model presented in the main text of the paper naturally encompasses various versions of the leapfrogging innovation game. We sketch here two such versions.

First, assume that innovation costs are only weakly increasing in the state, i.e. \(\chi_s \geq \chi_{s+1}\). This obviously encompasses the case, that we consider in the rest of the section, where innovation costs do not depend on the state, i.e. \(\chi_s = 1\) for all \(s\). Assume \(\tau_2 > 0\), then \(V^f_3 - V^f_2 > 0\) and thus \(V^f_3 - V^f_1 > 0\). Iterating the argument for \(s = 3, 4, \ldots\), we would conclude that \(V^f_{s+1} - V^f_s > 0\) and \(\lambda^f_s = 0\) for any \(s > 1\). Hence, followers never innovate at stages \(s > 1\).
also that now the value function $W$ in (34) equals $V^{f}_1$. This new version of the model assumes that, in order to become successful leaders, followers at any state $s > 1$ have to make two technological “leaps” to become leaders. First, a successful innovation propels them directly from $s > 1$ to $s = 1$. Then, a second successful innovations allows them to leapfrog the leader. In this setting, even when $\chi_s = 1$ for all $s$, the cumulative cost of leapfrogging the leader starting from any $s > 1$ is twice as large (it takes two successful innovations to leapfrog the leader) as the leapfrogging cost for $s = 1$. Following the same logic as the one outline above, it is straightforward to show again that if $\lambda^{f*}_1 > 0$ then necessarily $\lambda^{f*}_s = 0$ for all $s > 1$ (since $W = V^{f*}_1 = 0$ and $V^{f*}_s \geq V^{f*}_1$ then $W - V^{f*}_s = 0$) and therefore followers never innovate at states $s > 1$.

Second, it should be evident that the argument above can be pushed to the extreme by making $W = V_{s-1}$ for every $s > 1$. In this such version of the model it takes exactly $s$ successful innovations for a follower who is $s$ steps behind to finally leapfrog the leader. The cumulative leapfrogging cost at $s > 1$ is therefore exactly $s$ times the leapfrogging cost at $s = 1$. Once more, because of the linearity of the R&D technology, if $\lambda^{f*}_1 > 0$ then $\lambda^{f*}_s = 0$ for all $s > 1$.

It is important to note that the last two versions that we have outlined above are still better described as models of leapfrogging innovation, where a competitive fringe of followers faces higher leapfrogging costs the farther they are from the leader. In particular, these versions are not “step-by-step” models in the sense of Aghion et al. (2001), since they lack two central features of that strand of literature, namely a “neck-and-neck” state with low profits for all firms, and strictly increasing marginal costs of innovation that underpin a duopoly setting where the value of being a follower is strictly positive. For these reasons, the two versions of the leapfrogging model differ from both the “fast catch-up” and the “slow catch-up” cases in Acemoglu and Akcigit (2012). Likewise, the “leapfrogging and infringement” extension of Acemoglu and Akcigit (2012), where followers can also perform “frontier innovation” that allows them to leapfrog the leader by paying a state-contingent patent infringement fee, remains significantly different from our setup, as such extension does not assume free-entry and is conducted by means of numerical simulations for a small set of parameter values. As a result, we do not know analytically whether there exists a region of parameter values and polices under which multiple steady states exist and what the transitional dynamics looks like. All these are key contributions of our work.

**C. Policy reversal of lengthening patents’ duration**

In the main text, we discuss a policy of increasing markup ceiling, $m$, from an initial $H$ steady state, based on the calibrated model presented in Sect. 4.1. Here we conduct a similar policy by varying the parameter that governs the length of patent duration, $\tau$.

The economy is initial at an $H$ steady state under the same set of parameter as in Table 1 in period 1, where the patent duration is calibrated to be 20 periods (i.e. $1/\tau = 20$). The initial growth rate is 3.28%. In period 2, an unexpected and permanent
reduction of $\tau$ brings the patent duration to 22 periods (i.e. $\Delta 1/\tau = 2$) and at the same
time results in $m > \overline{M}(1/22)$. Panel (a) in Fig. 5 illustrates the dynamic path of the
economy following this policy. Clearly, after an initial outburst of growth, the economy
eventually evolves to steady state with low long-run growth at 2.83%.

Now suppose a policy maker who observes the tendency of the declining growth
decides to reverse the policy in period 15 by restoring the original $\tau$ (Panel (b) of
Fig. 5). As is clear from the figure, it does not stop the economy from continuing to
converge to an $L$ steady state, albeit the steady state growth is slightly improved from
before at 3.11%. One difference here from the policy experiment on $m$ we present in
Sect. 5.1 of the paper is that while the $L$ steady state growth $g^*_L$ is independent from
$m$, it increases in $\tau$ (see the proof of Corollary 1 in Appendix A.3). This is why as
we compare the $g^*_L$ associated with the policy $\tau = 1/22$ to the $g^*_L$ associated with
the policy reversal $\tau = 1/20$, the steady state growth improves. The policy reversal
mitigated partially the disastrous effect of the original policy on the long-run growth
of the economy.

To revert the equilibrium path to an $H$ steady state, however, requires a bigger policy
remedy. In Panel (c) of Fig. 5, we instead implement an unexpected and permanent
policy remedy in period 15 which reduces the length of the patent duration to 18
periods (e.g. $\Delta 1/\tau = -4$). Under this policy remedy, the condition that ensures the
only survival of the $H$ steady state holds, $m < \overline{M}(1/18)$. Therefore, the economy
eventually converges to the $H$ steady state associated with $\tau = 1/18$, which restores
the initial growth rate of 3.28%. The fact that we can restore the initial growth rate in
this example but not in the experiment on $m$ in Sect. 5.1 of the paper is that the growth
rate in an $H$ steady state, $g^*_H$ is independent of $\tau$ but increasing in $m$ (see the proof of
Corollary 1 in Appendix A.3).

The take-away from the policy experiment is two-fold. First, a small increase in
the length of patent protection from 20 to 22 periods when the economy is at a $H$
steady state can tip it over to converge to a $L$ steady state with large and negative
long-run consequence on growth, following a short-lived bout of growth immediately
following the policy. Second, to undo the small increase in the duration of patent
length and restore the original high growth rate, one requires a policy remedy that
over-compensates. In this particular case, shortening the length of patent duration to
18 periods will suffice.

D. The solution to planner’s problem

Using symmetry properties, we can drop the index $i$ and write the social planner’s
problem as

$$\max_{\Lambda_t, L_t} \int_0^\infty e^{-\rho t} \left( S_t \log \gamma + \log L_t \right)$$

s.t. $\dot{S}_t = \chi \Lambda_t$

$\Lambda_t \leq \overline{\Lambda}$

$\Lambda_t + L_t \leq 1$. 

\(\triangleright\) Springer
Fig. 5 Patent length policy and its reversals. Note: This figure illustrates how an economy parameterized according to Table 1 evolves when $\tau$ is reduced from $1/20$ to $1/22$ in period 2 (a); when $\tau$ is further restored to $1/20$ in period 15 (b); and when $\tau$ is instead increased to $1/18$ in period 15 (c). For a discussion, see Appendix C.
Assume that the constraint $\Lambda$ does not bind in equilibrium. Then the current value Hamiltonian is $H(S_t, \Lambda_t, \mu_t) \equiv S_t \log \gamma + \log(1 - \Lambda_t) + \mu_t \chi \Lambda_t$, where $\mu_t$ is the co-state variable. The optimality conditions are

$$\frac{\partial H(S_t, \Lambda_t, \mu_t)}{\partial \Lambda} = \frac{-1}{1 - \Lambda_t} + \mu_t \chi = 0$$
$$\frac{\partial H(S_t, \Lambda_t, \mu_t)}{\partial S} = \log \gamma = -\dot{\mu}_t + \rho \mu_t.$$

Differentiating the first equality with respect to time and combining it with the second equality yields

$$\frac{\dot{L}}{L} + \rho = \chi \log(\gamma) L.$$

In the steady state, $L^* = \frac{\rho}{\chi \log(\gamma)}$ and $\Lambda^* = 1 - \frac{\rho}{\chi \log(\gamma)}$. The consumption growth rate is $g_{SP} = \log(\gamma) \chi \Lambda^* = \chi \log(\gamma) - \rho$. We focus on the steady state comparisons, because in the social planner’s problem, the steady state is a source. The equilibrium path of $L_t$ in fact indicates that if $L_t > L^*$, then $L_t$ will increase without bound and if $L_t < L^*$, then $L_t$ will decrease without bound. In either case, $L_t$ will eventually violate the boundary conditions that $0 \leq L_t \leq 1$.

The consumption growth rate in the $H$ steady state in the decentralized economy is $g_{H}^* = \log(\gamma) \frac{(m-1)\chi - \rho}{m}$. This means that $g_{SP} > g_{H}^*$ if and only if

$$m < \log(\gamma) \left( \frac{\chi}{\rho} + 1 \right).$$

The consumption growth rate in the $L$ steady state in the decentralized economy is $g_{L}^* = \log(\gamma) \chi \frac{\tau \chi \Lambda}{\chi \Lambda - \rho}$. This means that $g_{SP} < g_{L}^*$ if and only if

$$\chi \log(\gamma) - \rho < \log(\gamma) \chi \frac{\tau \chi \Lambda}{\chi \Lambda - \rho} \iff \chi \log(\gamma) \left( \frac{(1 - \tau) \chi \Lambda - \rho}{\chi \Lambda - \rho} \right) < \rho.$$

When $0 < \tau < 1$, which is the relevant case, the condition becomes:

$$\tau > 1 - \frac{\rho}{\chi \Lambda} \left( \frac{\chi \Lambda - \rho}{\chi \log(\gamma)} + 1 \right).$$

These results are summarised in Proposition 3.
E. The model with constant intertemporal elasticity of substitution utility

In the baseline model, we assumed that households have a log period utility function, which amounts to assuming unit intertemporal elasticity of substitution. In this section, we relax this assumption by adopting a more general class of utility functions for households:

$$\int_0^\infty e^{-\rho t} \frac{C_t^{1-\sigma}}{1-\sigma} dt,$$

where $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution. All other elements of the model remain the same as in the baseline model. The consumption Euler equation becomes:

$$r_t = \rho + \sigma \frac{\dot{E}_t}{E_t} + (\sigma - 1) \log(\gamma) \dot{S}_t.$$  \hspace{1cm} (36)

In a steady state, the relationship between the interest rate and the rate of technological growth is now

$$r^* = \rho + (\sigma - 1) g^*.$$  

With a unit elasticity $\sigma = 1$ the steady state interest rate is equal to the rate of time preference $\rho$, as in the baseline model.

Under the more general utility function, the steady state interest rate depends positively (negatively) on growth when $\sigma$ is larger (smaller) than unity. This implies that when $\sigma > 1$, in a steady state with high technological growth and innovation, the interest rate will also be high, which would tend to self-confirm the high-growth situation by discouraging innovation by leaders. As discussed in Sect. 3.3, this is one of the sources of the multiplicity of steady states in our model. On the other hand, when $\sigma < 1$, high aggregate growth and innovation would lead to low interest rates, which would encourage leaders to innovate and would thus decrease the share industries in the contestable state. This by itself is a force that, in the long-run, would tend to push the economy toward a low-growth situation, playing against the self-confirmation of a high-growth steady state. This reasoning show that if $\sigma < 1$ is small enough, then the economy may not display in fact a multiplicity of steady states.

We can characterize analytically the structure of the steady states for a range of $\sigma$ in the following proposition and the proof follows.

**Proposition 4** There exist two constants $\sigma_H < 1 < \sigma_L$ such that for $\sigma \in (\sigma_H, \sigma_L)$ the economy has three steady states, $H$, $M$, and $L$. The $H$ and $L$ steady states are saddle path stable, while the $M$ steady state is unstable. Moreover, for $1 \leq \sigma < \sigma_L$, the three steady states are also ranked by their aggregate growth rates, $g^*_H > g^*_M > g^*_L$.

Fix an $m \in (\underline{M}, \overline{M})$ and all other parameters as in baseline model. There exists a range of $\sigma$, $(\sigma_H, \sigma_L)$ and $\sigma_H < 1 < \sigma_L$, in which there are three steady states. At
σ_H, the H and M steady states coincide where leaders become indifferent between innovating and not innovating. At σ_L, the M and L steady states coincide where the constraints on \( \Lambda \) becomes just binding. Figure 6 illustrates the three steady states, their extensive margins of innovation and growth rates, as we vary \( \sigma \). This figure is based on simulations of the model, keeping all parameters as in Table 1 and varying \( \sigma \) around unity. The red line denotes the steady states corresponding to a model with \( \sigma = 1 \) (i.e. the baseline model).

The figure shows that for low values of \( \sigma \) the only steady state that exists is the L steady state. For moderate values of \( \sigma \) around one, multiple steady states arise. However, when \( \sigma \) becomes too big, only H steady state survives. Beyond that point in fact the steady state interest rate is too high to warrant innovation by leaders.

**Proof of Proposition 4**

The representative household solves the following problem:

\[
\max \int_0^\infty e^{-\rho t} \frac{C_t^{1-\sigma}}{1-\sigma} \, dt
\]

\[\text{s.t.} \quad \int_0^\infty e^{-Rt} E_t \, dt \leq W(0),\]

where \( R_t \) is the compounded interest rate and \( E_t \) represents total spending at time \( t \):

\[R_t = \int_0^t r(\tau) d\tau,\]

\[E_t = \int_{[0,1]} p_i d_i dt.\]

The Cobb-Douglas form of the the consumption aggregate implies that the amount spent by the household on good \( i \) is the same across all products, giving

\[d_i = \frac{E_t}{p_{it}}.\]

Therefore, we can write the consumption aggregate as

\[\log C_t = \int_{[0,1]} \log \left( \frac{q_{it}}{p_{it}} E_t \right) d_i = \int_{[0,1]} \log \left( \frac{q_{it}}{p_{it}} \right) d_i + \log E_t \equiv \log Q_t + \log E_t,\]

where \( Q_t \) is proportional to the aggregate quality index and \( C_t = Q_t E_t \).

We can rewrite the consumer’s problem equivalently with a flow budget constraint,

\[\dot{a}_t = r_t a_t + I_t - E_t,\]

where \( a_t \) is the stock of savings (wealth) at time \( t \) and \( I_t \) is the total income (labor income and profit from firms) at time \( t \). We set up the current value Hamiltonian, \( \mathcal{H}(E_t, a_t, \mu_t) = \frac{(Q_t E_t)^{1-\sigma}}{1-\sigma} + \mu_t (r_t a_t + I_t - E_t). \) The first order conditions, \( \frac{\partial \mathcal{H}}{\partial E_t} = 0 \) and \( \frac{\partial \mathcal{H}}{\partial a_t} = \rho \mu_t - \dot{\mu}_t \), imply
Fig. 6 The model with the constant intertemporal elasticity of substitution preference. Note: This figure illustrates the structure of the steady states of the model extended to have constant intertemporal elasticity of substitution preference. In particular, it shows how the steady state extensive margin of innovation $\alpha^*$ and the steady state growth rate $g^*$ vary as we vary the elasticity of substitution parameter $\sigma$. For a discussion, see Appendix C.
\[ r_t = \rho + \sigma \frac{\dot{E}_t}{E_t} + (\sigma - 1) \frac{\dot{Q}_t}{Q_t} \]
\[ = \rho + \sigma \frac{\dot{E}_t}{E_t} + (\sigma - 1) \log(\gamma) \dot{S}_t, \]

which is (36). Also note that in the special case of log period utility \((\sigma = 1)\), we obtain the familiar \( \frac{\dot{E}_t}{E_t} = r_t - \rho \).

Let’s maintain all the parametric assumptions made in the baseline model and suppose \( m \in (\underline{M}, \bar{M}) \) so three steady states exist under the baseline assumption \( \sigma = 1 \). We characterize the structure of the steady states in this environment when \( \sigma \) deviates from 1.

The **H steady state** The highest feasible steady state value for \( \alpha^* \) is one, since in this case \( \lambda^* = 0 \). Let’s first assume that around the \( H \) steady state we have \( \lambda^f > 0 \). As usual, a steady state with high extensive margin will be associated with a low intensive margin \( \Lambda^* \). By taking \( \Lambda \) large enough, we can make sure that in a neighborhood of the steady state \( \Lambda^* = \lambda^f < \Lambda \), giving \( \omega_t = 1 \). Hence, the first order condition for \( \lambda^f \) in a neighborhood of a \( H \) steady state implies that

\[ V_1^f(t) = \frac{1}{\chi}. \]

The condition above implies that, in a neighborhood of the \( H \) steady state, \( \dot{V}_1^f = 0 \). Since \( \lambda^f = 0 \), a straightforward substitution in the definition of \( V_1^f \) gives

\[ \frac{r_t}{\chi} + \lambda^f = \Pi_t. \]

Combining the above equation with the facts that \( \Pi_t = (m - 1)L_t, \lambda^f = (1 - L_t)/\alpha_t \) and (36), we obtain

\[ \frac{\rho}{\chi} + \frac{\sigma}{\chi} \frac{\dot{L}}{L} + (\sigma - 1) \log(\gamma)(1 - L) + \frac{1 - L}{\alpha} = (m - 1)L. \]

Equation (40) defines the evolution of the economy around the \( H \) steady state, together with the condition

\[ \dot{\alpha} = \tau (1 - \alpha). \]

The \( H \) steady state is then characterized by

\[ \alpha^*_H = 1; \]
\[ L^*_H = \frac{1 + \rho/\chi + (\sigma - 1) \log(\gamma)}{m + (\sigma - 1) \log(\gamma)}; \]
\[ \lambda_{fH}^* = \frac{m - 1 - \rho/\chi}{m + (\sigma - 1) \log(\gamma)}; \]

\[ g_{H}^* = \log(\gamma) \frac{\chi(m - 1) - \rho}{m + (\sigma - 1) \log(\gamma)}. \]

Linearizing the system (40) and (41), we can show that the steady state is a saddle. We maintain the assumption from the baseline model that \( m > 1 + \rho/\chi \) to have a non-degenerate steady state. Moreover, \( \lambda_{fH}^* > 0 \) and \( L_{H}^* > 0 \) jointly requires \( \sigma > 1 - \frac{1 + \rho}{\log(\gamma)} \equiv \sigma_{H1}. \)

For the steady state to exist, we in addition require \( V_{l2}^* < 2/\chi \) so that leaders indeed do not have incentive to innovate. This gives us

\[ V_{l2}^* = \frac{(m - 1)L_{H}^* + \tau V_{l1}^*}{\rho + (\sigma - 1)g_{H}^* + \tau} \]

\[ = \frac{(m - 1)\chi \left[ 1 + \rho/\chi + (\sigma - 1) \log(\gamma) \right] + \tau (m + (\sigma - 1) \log(\gamma))}{(\rho + \tau)\chi (m + (\sigma - 1) \log(\gamma)) + (\sigma - 1) \log(\gamma) (\chi (m - 1) - \rho)\chi} < \frac{2}{\chi} \]

\[ \Rightarrow \sigma > 1 - \frac{\left( \frac{\rho + \chi}{\chi - \rho - \tau} - m \right) (\chi - \rho - \tau)}{\log(\gamma) (\chi (m - 1) + \tau)} \]

\[ = 1 - \frac{(M - m) (\chi - \rho - \tau)}{\log(\gamma) (\chi (m - 1) + \tau)} \equiv \sigma_{H2}. \]

We maintain the assumption from the baseline that \( \chi > \rho + \tau \). For the set of parameters under which the steady state exists in the baseline model (i.e. \( m < \overline{M} \)), the steady state exists in this extended model as long as \( \sigma \) is greater than \( \sigma_{H2} \), which is a number less than 1.

To compare \( \overline{\sigma}_{H1} \) and \( \overline{\sigma}_{H2} \), we first note that

\[ \frac{\rho + \chi}{\chi} - \frac{(\overline{M} - m) (\chi - \rho - \tau)}{\chi (m - 1) + \tau} \]

\[ = \frac{(2\chi - \tau)\chi \left( m - 1 - \frac{\rho}{\chi} \right)}{\chi (\chi (m - 1) + \tau)}, \]

which has the same sign as \( m - 1 - \frac{\rho}{\chi} \). Recall that \( m > \overline{M} = 1 + \frac{(2\rho + \tau)(\chi N - \rho)}{\chi ((\chi - \tau)\Lambda - \rho)} \) and \( \overline{M} \) is decreasing in \( \Lambda \). This implies

\[ m > \overline{M} > \lim_{\Lambda \to +\infty} \overline{M} = 1 + \frac{2\rho + \tau}{\chi - \tau} > 1 + \rho/\chi. \]
Therefore, \( m - 1 - \frac{\rho}{\chi} > 0 \), \( \frac{\rho + \chi}{\chi} > \frac{(M-m)(\chi-\rho-\tau)}{\chi(m-1)+\tau} \), and in turn

\[ \sigma_{H1} < \sigma_{H2}. \]

Define \( \sigma_H = \sigma_{H2} \). The \( H \) steady state exists as long as \( \sigma \geq \sigma_H \). At \( \sigma_H \), leaders are indifferent between innovating or not. For a \( \sigma \) that is infinitesimally smaller than \( \sigma_H \), leaders will have strictly prefer to innovate.

The \( M \) and \( L \) steady states In the \( M \) and \( L \) steady states, both leaders and followers innovate at the contestable state. The first order conditions for \( \lambda^l \) and \( \lambda^f \) give \( V^l_2(t) = 2\omega_t/\chi = 2V^l_1(t) \). Substituting these conditions into the value functions and after appropriate calculations we obtain the two equations:

\[
\begin{align*}
\Pi &= (2\lambda^f - \frac{\tau}{\chi})\omega \\
\dot{\omega} &= (r + \tau - \chi\lambda^f)\omega. 
\end{align*}
\] (42)

From the first equation in (42) and \( \Pi = (m - 1)L \) we can solve out \( L \):

\[ L = \frac{2\lambda^f - \tau/\chi}{m - 1}\omega, \]

which, together with the labor market clearing condition, implies

\[ 1 - \frac{2\lambda^f - \tau/\chi}{m - 1}\omega = \alpha(\lambda^f + \lambda^l). \] (43)

From the second equation in (42), in the steady state \( r + \tau - \chi\lambda^f = 0 \). Combined with the Euler equation derived at the beginning of this section, we have

\[ \rho + (\sigma - 1)g + \tau - \chi\lambda^f = 0. \] (44)

The difference between an \( M \) and a \( L \) steady state is that in an \( M \) steady state, the skilled labor supply does not bind (\( \lambda^l + \lambda^f < \Lambda \)) and the skill premium is one (\( \omega = 1 \)), whereas in a \( L \) steady state, the opposite is true: \( \lambda^l + \lambda^f = \Lambda \) and \( \omega > 1 \).

Then we can use four equations to characterize an \( M \) steady state

\[
\begin{align*}
\alpha &= \frac{\tau}{r+\chi\lambda^l} \\
g &= \log(y)\chi\alpha(\lambda^f + \lambda^l) \\
\rho + (\sigma - 1)g + \tau - \chi\lambda^f &= 0 \\
1 - \frac{2\lambda^f - \tau/\chi}{m - 1} &= \alpha(\lambda^f + \lambda^l) 
\end{align*}
\] (45)

The first equation becomes from the evolution of the extensive margin \( \alpha_t \). The second equation is the definition of the growth rate. The third equation is (44). The last equation is (43), where \( \omega = 1 \) in an \( M \) steady state. From these four equations, we can solve for the \( M \) steady state endogenous variables: \( \alpha^*_M, g^*_M, \lambda^f_M \) and \( \lambda^l_M \).
We can use another set of four equations to characterize a $L$ steady state

$$\begin{align*}
\alpha &= \frac{\tau}{\tau + \chi(\lambda - \lambda^f)} \\
g &= \log(\gamma) \chi \alpha \Lambda \\
\rho + (\sigma - 1)g + \tau - \chi \lambda^f &= 0 \\
1 - \frac{2\lambda^f - \tau / \chi}{m - 1} \omega &= \alpha \Lambda
\end{align*} \tag{46}$$

From these four equations, we can solve for the $L$ steady state endogenous variables: $\alpha_L^*, g_L^*, \lambda_f^* L$, and $\omega_L^*$.

Let’s focus on the $M$ steady state first. Combining the second and fourth equation in (45), we have one equation that links $g$ to $\lambda^f$:

$$g = \log(\gamma) \chi \left[ m - 1 - 2\lambda^f + \tau / \chi \right]. \tag{47}$$

Together with the third equation in (45), we can solve out the $M$ steady state explicitly

$$\begin{align*}
\alpha_M^* &= \frac{(\chi - \tau)(m - 1) - (2\rho + \tau) - 2\tau(\sigma - 1) \log(\gamma)}{\rho(m - 1) - [\tau - \chi(m - 1)](\sigma - 1) \log(\gamma)}; \\
L_M^* &= \frac{2\rho + \tau + 2\chi(\sigma - 1) \log(\gamma)}{\chi(m - 1) + 2\chi(\sigma - 1) \log(\gamma)}; \\
\lambda_f^* M &= \frac{(\sigma - 1) \log(\gamma) \chi [m - 1 + \tau / \chi] + (m - 1)(\rho + \tau)}{[m - 1 + 2(\sigma - 1) \log(\gamma)] \chi}; \\
g_M^* &= \log(\gamma) \frac{\chi(m - 1) - 2\rho - \tau}{m - 1 + 2(\sigma - 1) \log(\gamma)}.
\end{align*}$$

The $\lambda_f^* M$ is implied in the last equation of (45). Rearranging terms,

$$\lambda_f^* M = \frac{\tau \left( (m + 1)\lambda^f M - m + 1 - \tau / \chi \right)}{(m - 1)(\chi - \tau) + \tau - 2\chi \lambda^f M^*}, \tag{48}$$

which is increasing in $\lambda_f^* M$. All these endogenous variables are well-defined when $\sigma = 1$. Let’s differentiate $\lambda_f^* M$ with respect to $\sigma - 1$.

$$\frac{d\lambda_f^* M}{d(\sigma - 1)} = \frac{\log(\gamma)(m - 1)\chi [(m - 1)\chi - 2\rho - \tau]}{[m - 1 + 2(\sigma - 1) \log(\gamma)]^2 \chi^2}.$$ 

Recall that $m > M = 1 + \frac{(2\rho + \tau)(\chi - \rho)}{\chi((\chi - \tau)\Lambda - \rho)} > 1 + \frac{2\rho + \tau}{\chi}$. Hence,

$$\frac{d\lambda_f^* M}{d(\sigma - 1)} > 0.$$
This means, as $\sigma$ decreases below 1, both $\lambda^f_M$ and $\lambda^l_M$ will decrease until $\lambda^l_M$ becomes zero, at which point the $M$ steady state coincides with the $H$ steady state when leaders are indifferent between innovating and not innovating. To see this point, when $\lambda^l_M = 0$, from (48), $\lambda^f_M$ becomes

$$\lambda^f_M = \frac{m - 1 + \tau/\chi}{m + 1}.$$ 

Evaluate the $\lambda^f_H$ at $\sigma = \sigma_H$:

$$\lambda^f_H = \frac{m - 1 - \rho/\chi}{m - \frac{(M-m)(\chi-\rho-\tau)}{\chi(m-1)+\tau}} = \frac{m - 1 + \tau/\chi}{m + 1} = \lambda^f_M.$$ 

This also means, as $\sigma$ rises above 1, both $\lambda^f_M$ and $\lambda^l_M$ will increase until the sum hits the fixed supply: $\lambda^f_M + \lambda^l_M = \overline{\lambda}$. At this point, as we will show below, the $M$ steady state coincides with the $L$ steady state where $\omega = 1$.

Let’s focus on the $L$ steady state now. Combining the first two equations in (46) and cancelling out $\alpha$, we have

$$\frac{g}{\log(\gamma)\chi\overline{\lambda}} = \frac{\tau}{\tau + \chi(\overline{\lambda} - \lambda^f)}.$$ 

Combining the above with the third equation in (46), we can infer the $L$ steady state $\lambda^f_L$ from

$$\chi^2\lambda^f_L^2 - (\rho + 2\tau + \chi\overline{\lambda})\chi\lambda^f_L + (\rho + \tau)(\tau + \chi\overline{\lambda}) + (\sigma - 1)\log(\gamma)\tau\chi\overline{\lambda} = 0. \quad (49)$$

Under our assumption of $m \in (M, \overline{M})$, we know when $\sigma = 1$ there exists a well-defined $L$ steady state. When $\sigma = 1$, the above quadratic has two roots: $\lambda^f = \frac{\rho + \tau}{\chi}$ and $\lambda^f = \overline{\lambda} + \frac{\tau}{\chi}$ (omitted because it is greater than $\overline{\lambda}$). The smaller root is the R&D intensity of the followers in the $L$ steady state in the baseline model, $\lambda^f_L^*$, and we also know in that steady state $\omega^*_L > 1$. As $\sigma$ increases above 1, the quadratic function shifts up and the smaller root, $\lambda^f_L^*$, increases, which in turn implies that $\alpha^*_L$ increases (see the first equation of (46)). Now from the fourth equation in (46), we deduce that the steady state $\omega^*_L$ must decrease. Therefore, as $\sigma$ increases, the smaller root to (49) defines the $L$ steady state level of $\lambda^f$ until the implied $\omega^*_L$ decreases to 1, at which point the $L$ steady state coincides with the $M$ steady state where the constraint on skilled labor supply becomes just binding.

To see this point, note how the solution to (45) when $\lambda^f + \lambda^l = \overline{\lambda}$ must also solve (46) when $\omega = 1$ and vice versa.

Finally, we show that the larger root of this quadratic equation (49) can never be a $L$ steady state. Since $\sigma$ only shifts the quadratic function up and down, the larger root
will always be strictly larger than the $\lambda_{L}^{f*}$ when $L$ and $M$ steady states coincide as we discuss above. Suppose the larger root, $\lambda_2^f$ occurs in a $L$ steady state. Then in that steady state, the extensive margin $\alpha$ must be larger than the extensive margin when $L$ and $M$ steady states coincide. This also means, $\omega$ in that steady state must be strictly smaller than the skill premium when $L$ and $M$ steady states coincide, which we know is 1. This contradicts the definition of a $L$ steady state.

Let $\sigma_L$ be the $\sigma$ at which the $L$ and $M$ steady states coincide and let $\sigma_L$ be the $\sigma$ when the smaller root of the quadratic equation is $\frac{\tau}{2\chi}$ and $\sigma_L < 1$. We have shown that for $\sigma \in (\sigma_L, \sigma_H)$, the $L$ steady state exists. $\lambda_{L}^{f*}$ is given by the smaller root of equation (49) and the other steady state variables can be derived by

\[
\alpha_{L}^{*} = \frac{\tau}{\tau + \chi(\Lambda - \lambda_{L}^{f*})}; \quad g_{L}^{*} = \alpha_{L}^{*} \log(y) \chi \Lambda; \quad L_{L}^{*} = 1 - \alpha_{L}^{*} \Lambda.
\]

We next show that $\sigma_L < \sigma_H$, such that the $L$ steady state is defined whenever the $H$ steady state is defined and $\sigma < 1$. Substituting $\lambda_f$ with $\frac{\tau}{2\chi}$ in (49), we can rearrange to obtain

\[
(1 - \sigma_L) \log(\gamma) = \frac{(2\rho + \tau)(2\chi \Lambda + \tau)}{4\tau \chi \Lambda}.
\]

Recall $\sigma_H = \sigma_{H2}$ and

\[
(1 - \sigma_H) \log(\gamma) = \frac{\chi + \rho - m(\chi - \rho - \tau)}{\chi(m - 1) + \tau}.
\]

We can derive the following inequalities

\[
(1 - \sigma_H) \log(\gamma) < \frac{(2\rho + \tau)\rho(\chi \Lambda - \tau - \rho)}{2\chi(\rho + \tau)(\chi \Lambda - \rho) - \tau^2 \Lambda \chi} = \frac{2\rho + \tau}{\chi \Lambda} \frac{\rho(\chi \Lambda - \tau - \rho)}{2(\rho + \tau) (\chi - \frac{\rho}{\Lambda}) - \tau^2} \leq \frac{2\rho + \tau}{\chi \Lambda} \frac{\rho(\chi \Lambda - \tau - \rho)}{2(\chi - \tau)}.
\]

The first inequality is obtained by replacing $m$ by $M$ since $(1 - \sigma_H) \log(\gamma)$ decreases in $m$ and $m > M$. The second inequality is obtained by replacing $\left(\chi - \frac{\rho}{\Lambda}\right)$ on the denominator by $\frac{\chi \tau}{\rho + \tau}$ since $\Lambda > \frac{\rho + \tau}{\chi}$. Now, we have
(1 − σ_H) \log(γ) < (1 − σ_L) \log(γ)
\iff \frac{\rho(\chi \Lambda − τ − \rho)}{2\chi − τ} < \frac{2\chi \Lambda + τ}{4}
\iff 4\rho(\chi \Lambda − τ − \rho) < (2\chi − τ)(2\chi \Lambda + τ)
\iff 2\chi(2\rho − 2\chi + τ)\Lambda < (2\chi − τ + 4\rho(\rho + τ)),

which is always true. Because we maintain the assumption that \chi > \rho + \tau, the left hand side of the above inequality is negative whereas the right hand side is positive. Hence, we conclude

\sigma_L < \sigma_H < 1.

This means, the L steady state is always defined for any \sigma < 1 under which the H steady state is also defined.

The above discussion can be summarized in the first part of Proposition 4. We now show how aggregate growth is ordered in the three steady states as described in the proposition. First, we show that the growth rate in the H steady state is always higher than that in the M steady state. Since \(g^*_i = \log(γ)χ(1 − L^*_i)\) for \(i = M, H\), it suffices to show that \(L^*_H < L^*_M\).

\[L^*_H < L^*_M \iff \frac{1 + \rho/\chi + (\sigma - 1) \log(γ)}{m + (\sigma - 1) \log(γ)} < \frac{2\rho + \tau + 2\chi(\sigma - 1) \log(γ)}{\chi(m - 1) + 2\chi(\sigma - 1) \log(γ)}\]

\[\iff (\sigma - 1) \log(γ) > -\frac{\chi + \rho - m(\chi - \rho - \tau)}{\chi(m - 1) + \tau}\]

\[\iff \sigma > \sigma_H,\]

an assumption made in Proposition 4. Hence, we have \(g^*_H > g^*_M\).

Next we order \(g^*_M\) and \(g^*_L\). We first introduce the following Lemma which can be proved by contradictions.

**Lemma 1** In the M and L steady states, we have \(\lambda^*_M \geq \lambda^*_L\), if \(\sigma \geq 1\).

**Proof** The case of \(\sigma = 1\) is discussed in the baseline model, in which case \(\lambda^*_M = \lambda^*_L\). Suppose \(\sigma > 1\) and we prove by contradiction. Suppose \(\lambda^*_M \leq \lambda^*_L\). Since \(\lambda^*_i = \frac{r^*_i + \tau}{\chi}\), for \(i = M, L\), it implies that \(r^*_M \leq r^*_L\). Since \(r^*_i = \rho + (\sigma - 1)g^*_i\), it implies that \(g^*_M \leq g^*_L\). Since \(g^*_i = \log(γ)χ(1 − L^*_i)\), we have \(L^*_M \geq L^*_L\). On the other hand, from the first equation of (42), it must be true that

\[L^*_M = \frac{2\lambda^*_M - \tau}{m - 1} < \frac{2\lambda^*_L - \tau}{m - 1} < \frac{2\lambda^*_L - \tau}{m - 1} \omega_L = L^*_L,\]

since \(\omega_L > 1\). We reach a contradiction. \(\Box\)
This, together with the equation $\rho + (\sigma - 1)g_i^{*} + \tau - \chi \lambda_i^{f*} = 0$ for $i = M, L$, implies that $g_M^{*} > g_L^{*}$ as long as $\sigma > 1$. This concludes the proof for $g_H^{*} > g_M^{*} > g_L^{*}$ for $1 \leq \sigma < \sigma_L$. The proof of the local stability properties of the steady states is available upon request.

## F. The model with a quadratic cost of innovation

In this section we modify the model to introduce a quadratic cost of innovation to both leaders and followers. Suppose the cost of innovation is the following:

$$\phi_j \lambda + \frac{1}{2} \theta_j \lambda^2, \quad j = 1, 2,$$

where $j = 1$ is for leaders and $j = 2$ is for followers and $\phi_j, \xi_j > 0$ are parameters of the model. Since this effectively imposes decreasing return on innovation at the firm level, we then abandon the assumption of maximum supply $\Lambda$ of skilled labor. We solve and simulate this modified model and examine if the baseline key properties of the steady states survive these modifications.

In Fig. 7, we plot the steady state values of the extensive margin of innovation, $\alpha^*$, and the growth rate, $g^*$, against different values of the markup $m$ from the modified model. Comparing this figure to Fig. 1 from the baseline model, we confirm that the structure of the steady states under the quadratic cost of innovation remains similar to that in the linear model.

The mathematical derivations of the steady states in the model with quadratic costs are as follows.

Replace the linear cost of innovation in the baseline model with the following quadratic cost. In order to achieve an arrival rate of innovation of $\lambda$ (to reduce notation we normalize $\xi = 1$), the firm needs to employ the following amount of skilled labor:

$$\phi_j \lambda + \frac{1}{2} \theta_j \lambda^2,$$

where $\phi_j$ and $\xi_j$ are parameters of the cost function for leaders ($j = 1$) and followers ($j = 2$). The value functions of leaders and followers are given as follows.

$$rV_1^l = \max_{\lambda l \geq 0} \lambda l - \phi_1 \lambda l^2 + \lambda l (V_2^l - V_1^l) + \lambda l (V_1^f - V_1^l) + \dot{V}_1^l$$

(52)

$$rV_1^f = \max_{\lambda l \geq 0} \phi_2 \lambda f^2 + \lambda f (V_1^f - V_1^l) + \dot{V}_1^f$$

(53)

The FOCs imply
Fig. 7  The model with quadratic costs of innovation. Note: This figure illustrates the structure of the steady states of the model extended to have quadratic cost of innovation. In particular, it shows how the steady state extensive margin of innovation $\alpha^*$ and the steady state growth rate $g^*$ vary as we vary the markup parameter $m$. For a discussion, see Appendix 6.
Both leaders and followers innovating

Focus on the steady states where both leaders and followers innovate. Subtracting (52) from (50) and rearranging,

$$(r + \tau + \lambda^l)(V^l_2 - V^l_1) = \phi_1 \lambda^l + \frac{1}{2} \xi_1 \lambda^{l2} + \lambda^f (V^l_1 - V^f_1),$$

where $V^l_2 - V^l_1$ is given by (54) and $V^l_1 - V^f_1$ is given by (55), and $r = \rho$ in a steady state. This implies the first equation that involves $\lambda^f$ and $\lambda^l$:

$$(\rho + \tau)(\phi_1 + \xi_1 \lambda^l) + \frac{1}{2} \xi_1 \lambda^{l2} = \lambda^f (\phi_2 + \xi_2 \lambda^f).$$ (56)

Subtracting (53) from (51) and rearranging,

$$V^f_2 - V^f_1 = \frac{\phi_2 \lambda^f + \frac{1}{2} \xi_2 \lambda^{f2} - \lambda^f (V^l_1 - V^f_1)}{r + \tau + \lambda^l} = \frac{-\frac{1}{2} \xi_2 \lambda^{f2}}{r + \tau + \lambda^l},$$

where the last equality follows from substituting $V^l_1 - V^f_1$ by (55).

Subtracting (53) from (52) and rearranging,

$$(r + 2\lambda^f)(V^l_1 - V^f_1) = \Pi^* - \phi_1 \lambda^l - \frac{1}{2} \xi_1 \lambda^{l2} + \lambda^l (V^l_2 - V^l_1) + \phi_2 \lambda^f + \frac{1}{2} \xi_2 \lambda^{f2} - \lambda^l (V^f_2 - V^f_1),$$

where $V^l_1 - V^f_1$ is given by (55), $V^l_2 - V^l_1$ is given by (54), $V^f_2 - V^f_1$ is given by (57). Substituting these terms in the above equation, we have

$$(r + 2\lambda^f)(\phi_2 + \xi_2 \lambda^f) = \Pi^* + \frac{1}{2} \xi_1 \lambda^{l2} + \lambda^l \left( (r + \tau)(\phi_2 + \frac{1}{2} \xi_2 \lambda^f) + \lambda^l (\phi_2 + \xi_2 \lambda^f) \right)$$

$$\frac{r + \tau + \lambda^l}{r + \tau + \lambda^l}.$$ (58)

Note that in a steady state where both leaders and followers innovate, the extensive margin is given by

$$\alpha^* = \frac{\tau}{\tau + \lambda^l}.$$
\[ \Pi^* = (m - 1) \left[ 1 - \alpha^* \left( \phi_1 \lambda^l + \frac{1}{2} \xi_1 \lambda^{l2} + \phi_2 \lambda^f + \frac{1}{2} \xi_2 \lambda^{f2} \right) \right] \]

\[ = (m - 1) \left[ 1 - \frac{\tau}{\tau + \lambda^l} \left( \phi_1 \lambda^l + \frac{1}{2} \xi_1 \lambda^{l2} + \phi_2 \lambda^f + \frac{1}{2} \xi_2 \lambda^{f2} \right) \right], \]

which we can plug in (58) together with \( r = \rho \) to obtain

\[ \xi_2 \left[ \frac{1}{2} \left( \frac{\rho + \tau + 2 \lambda^l}{\rho + \tau + \lambda^l} - \frac{(m - 1) \tau}{\tau + \lambda^l} \right) - 2 \right] \lambda^f - \left( \frac{(m - 1) \phi_2 \tau}{\tau + \lambda^l} + \phi_2 \rho \xi_2 \right) \lambda^f \]

\[ + (m - 1) \left[ 1 - \frac{\tau}{\tau + \lambda^f} \left( \phi_1 \lambda^f + \frac{1}{2} \xi_1 \lambda^{f2} \right) \right] + \frac{1}{2} \xi_1 \lambda^{l2} - \rho \phi_2 = 0. \]  

(59)

Equations (56) and (59) form a system of equations, from which we can solve for \( \lambda^l \) and \( \lambda^f \), which give us the steady state \( \lambda^{l*} \) and \( \lambda^{f*} \).

Only followers innovating  
Now consider the steady state, where only followers innovate. In this steady state, \( \alpha^* = 1 \) and \( \lambda^{l*} = 0 \).

The value functions, (52) and (53), at the steady state become

\[ rV_1^l = \Pi + \lambda^f (V_1^f - V_1^l) \]

\[ rV_1^f = -\phi_2 \lambda^f - \frac{1}{2} \xi_2 \lambda^{f2} + \lambda^f (V_1^l - V_1^f). \]

Taking the difference of the above two equations, we have

\[ (r + 2 \lambda^f ) (V_1^l - V_1^f ) = \Pi + \phi_2 \lambda^f + \frac{1}{2} \xi_2 \lambda^{f2}. \]  

(60)

Note that the profit is given by

\[ \Pi = (m - 1) L = (m - 1) \left( 1 - \phi_2 \lambda^f - \frac{1}{2} \xi_2 \lambda^{f2} \right). \]  

(61)

Plugging (55) and (61) in (60) and replace \( r \) by the steady state value \( \rho \), we have

\[ (\rho + 2 \lambda^f ) (\phi_2 + \xi_2 \lambda^f ) = (m - 1) \left( 1 - \phi_2 \lambda^f - \frac{1}{2} \xi_2 \lambda^{f2} \right) + \phi_2 \lambda^f + \frac{1}{2} \xi_2 \lambda^{f2} \]

\[ \Rightarrow \xi_2 \left( \frac{1}{2} m + 1 \right) \lambda^{f2} + (\rho \xi_2 + m \phi_2) \lambda^f + \rho \phi_2 - m + 1 = 0. \]

from which we can solve for the steady state value for \( \lambda^f \), \( \lambda^{f*} \).

From the corner solution for \( \lambda^f \), we can infer that

\[ V_2^l - V_1^l < \phi_1. \]
Taking the difference of $V^l_2$ and $V^l_1$, we have
\[ V^l_2 - V^l_1 = \frac{\lambda^f_*}{\rho + \tau} (V^l_1 - V^f_1). \]

From (60), we derive
\[ V^l_2 - V^l_1 = \frac{\lambda^f_*}{\rho + \tau} \left( \frac{\Pi + \phi_2 \lambda^f_* + \frac{1}{2} \xi_2 \lambda^f_*^2}{\rho + 2 \lambda^f_*} \right) = \frac{\lambda^f_*}{\rho + \tau} \left( (m - 1) - (m - 2) \left( \phi_2 \lambda^f_* + \frac{1}{2} \xi_2 \lambda^f_*^2 \right) \right) < \phi_1. \]

This is the condition for the existence of the steady state where leaders indeed do not innovate.

\[ \text{References} \]

Acemoglu, D., Akcigit, U.: Intellectual property rights policy, competition and innovation. J. Eur. Econ. Assoc. 10(1), 1–42 (2012)

Aghion, P., Howitt, P.: A model of growth through creative destruction. Econometrica 60(2), 323–351 (1992)

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., Li, H.: A theory of falling growth and rising rents. Rev. Econ. Stud. (forthcoming) (2022)

Aghion, P., Harris, C., Howitt, P., Vickers, J.: Competition, imitation, and growth with step-by-step innovation. Rev. Econ. Stud. 8, 467–492 (2001)

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P.: Competition and innovation: an inverted-U relationship. Q. J. Econ. 120(2), 701–728 (2005)

Atkeson, A., Burstein, A.: Aggregate implications of innovation policy. J. Polit. Econ. 127(6), 2625–2683 (2019)

Barro, R. J., Sala-i-Martin, X.: Economic Growth, McGraw-Hill (1995)

Boldrin, M., Levine, D.K.: Against Intellectual Monopoly. Cambridge University Press, Cambridge (2008)

Budd, C., Harris, C., Vickers, J.: A model of the evolution of duopoly: does the asymmetry between firms tend to increase or decrease? Rev. Econ. Stud. 60(3), 543–573 (1993)

Chu, A.C., Furukawa, Y., Mallick, S., Peretto, P., Wang, X.: Dynamic effects of patent policy on innovation and inequality in a Schumpeterian economy. Econ. Theor. 71, 1429–1465 (2021)

Covarrubias, M., Gutiérrez, G., Philippon, T.: From Good to Bad Concentration? U.S. Industries over the Past 30 Years,” NBER Macroeconomics Annual Conference (2019)

Denicòli, V.: Growth with non-drastic innovations and the persistence of leadership. Eur. Econ. Rev. 45, 1399–1413 (2001)

Dinopoulos, E., Syropoulos, C.: Rent protection as a barrier to innovation and growth. Econ. Theor. 32, 309–332 (2007)

Etro, F.: Innovation by leaders. Econ. J. 114, 281–303 (2004)

García-Maciá, D., Hsieh, C.-T., Klenow, P.: How destructive is innovation? Econometrica 87(5), 1507–1541 (2019)

Gersbach, H., Schmutzler, A.: Endogenous spillovers and incentives to innovate. Econ. Theor. 21, 59–79 (2003)

Gilbert, R.T., Newbery, D.M.G.: Preemptive patenting and the persistence of monopoly. Am. Econ. Rev. 72(3), 514–526 (1982)

Grossman, G.M., Helpman, E.: Quality ladders in the theory of growth. Rev. Econ. Stud. 58(1), 43–61 (1991)
Haltiwanger, J.: Top Ten Signs of Declining Business Dynamism and Entrepreneurship in the U.S., Kauffman Foundation New Entrepreneurial Growth Conference (2015)

Harris, C., Vickers, J.: Racing with uncertainty. Rev. Econ. Stud. 54(1), 1–21 (1987)

Hörner, J.: A perpetual race to stay ahead. Rev. Econ. Stud. 71, 1065–1088 (2004)

Kiedaisch, C.: Growth and welfare effects of intellectual property rights when consumers differ in income. Econ. Theor. 72, 1121–1170 (2021)

Klette, T.J., Kortum, S.: Innovating firms and aggregate innovation. J. Polit. Econ. 112(5), 986–1018 (2004)

Liu, E., Mian, A., Sufi, A.: Low interest rates, market power and productivity growth. Econometrica 90(1), 193–221 (2022)

Loecker, J.D., Eeckhout, J., Unger, G.: The rise of market power and the macroeconomic implications. Q. J. Econ. 135(2), 561–644 (2020)

Romer, P.: Endogenous technological change. J. Polit. Econ. 98, 71–102 (1990)

Segerstrom, P.S.: Intel economics. Int. Econ. Rev. 48(1), 247–280 (2007)

Segerstrom, P.S., Zolnierek, J.M.: The R&D incentives of industry leaders. Int. Econ. Rev. 40(3), 745–766 (1999)

Stein, J.C.: Waves of creative destruction: firm-specific learning-by-doing and the dynamics of innovation. Rev. Econ. Stud. 64, 265–288 (1997)

Suzuki, K.: Corporate tax cuts in a Schumpeterian growth model with an endogenous market structure. J. Public Econ. Theory 24, 324–347 (2021)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.