# Innovation, Speculation and Competitive Growth*

Vincenzo Denicolò
University of Bologna and University of Leicester

Piercarlo Zanchettin
University of Leicester

June 4, 2011

**Abstract**

In a classic paper, Hirshleifer (1971) noted that technological change affects the prices of various assets traded in the economy. He then argued that innovators could be rewarded by the speculative profits they can obtain thanks to their inside information as to the occurrence of innovations. We propose a tractable model of endogenous growth that formalises this insight. We then use the model to assess two conjectures made by Hirshleifer: (i) that speculative profits can generate excessive investment in R&D when they add to monopoly rents guaranteed by patent protection (the weak overinvestment hypothesis); (ii) that speculative profits can lead to excessive investment in R&D even in a perfectly competitive economy (the strong overinvestment hypothesis). The analysis confirms the weak overinvestment hypothesis, but casts doubts on the strong one.

*We are grateful to Gianni De Fraja, Ludovic Renou and seminar participants in Paris for useful comments.

# 1 Introduction

Eli Whitney, the inventor of the cotton gin, spent many years seeking patent protection and lost most of his profits trying to enforce it. After coming close to bankruptcy, he eventually turned to other projects. Unsuccessful as it may have been, his business model accords with the tenet of current theories of endogenous growth – that inventors can only be rewarded by monopoly rents guaranteed by patents.

However, both Eli Whitney and those theories may have overlooked other sources of reward. As noted by Hirshleifer (1971), technological change generally affects the equilibrium price of various assets traded in the economy. Since for a time the inventor has access to superior knowledge as to the occurrence of his own innovation, he can reap speculative profits by engaging in insider trading:

> The cotton gin had obvious speculative implications for the price of cotton, the value of slaves and of cotton-bearing land, the site value of key points in the transport network that sprang up. There were also predictable implications for competitor industries (wool) and complementary ones (textiles, machinery). It seems very likely that some forethoughted individuals reaped speculative gains on these developments, though apparently Whitney did not. And yet, he was the first in the know, the possessor of an unparalleled opportunity for speculative profit. (p. 571)

Hirshleifer recognised that inventors can capture only a fraction of the pecuniary effects of innovations, but claimed that if that fraction is large enough, speculative profits "may more than suffice for an appropriate inducement to invention." This paper provides what is, to the best of our knowledge, the first formalisation of Hirshleifer's insight. It develops a tractable model of endogenous growth in a perfectly competitive economy where inventive activity can be sustained by speculative gains rather than monopoly rents.

Hirshleifer actually claimed that speculative profits can not only sustain innovation, but also lead to overinvestment in research. We distinguish between a weak and a strong version of this claim, both advanced by Hirshleifer. The *weak overinvestment hypothesis* is that overinvestment in R&D can occur when speculative profits add to monopoly rents guaranteed by patent protection. After noting that a perfectly discriminating patent holder can capture the entire social value of the innovation even without speculating, Hirshleifer concluded that

> [...] the perfectly discriminating patent holder [...] is in a position to reap speculative profits, too; counting these as well, he would clearly be overcompensated. (p. 572)

The *strong overinvestment hypothesis* claims that speculative profits by themselves can lead to overinvestment in research:

> there is no logically necessary tie between the size of the technological benefit on the one hand, and the amplitude of the price shifts that create speculative opportunities on the other. [...] A relatively minor shift in locomotive technology, for example, might lead railroad planners to select an entirely different route for a new line, with drastic upward and downward shifts of land values. (p. 572)

Our analysis confirms the weak overinvestment hypothesis but casts doubts on the strong one. The intuitive reason why the validity of the strong hypothesis is dubious, even assuming that inventors can fully appropriate the pecuniary effects of innovations, is that there do exist a relation between speculative opportunities and the size of innovations after all. For example, if a minor shift in locomotive technology induces railroad planners to select an alternative route, then the two routes must be close substitutes both before and after the shift. If this is so, then changes in land value cannot be ample. In our formal model, tradeable assets appreciate only to the extent that innovations increase their productivity. The social value of innovations, by contrast, is the total increase in factor productivity, which includes also the increase in the productivity of labour. Because labour is non alienable, and hence cannot be the object of speculation, inventors rewarded by speculative gains only are necessarily undercompensated.

The validity of the weak hypothesis is also less obvious than it might seem. The problem here is that patent protection crowds out speculative gains. When the patent holder obtains the entire technological benefit from the innovation thanks to patent protection, asset prices do not change at all. Thus, with full patent protection the opportunities for speculation vanish. However, we show that overinvestment in R&D can occur when patent protection is appropriately limited. This is possible because the total incentive to innovate (which includes both monopoly rents and speculative gains) is not necessarily increasing in the level of patent protection.

Other papers have argued that innovation can be sustained in a perfectly competitive economy with no monopoly rents. Hellwig and Irmen (2001), building on Bester and Petrakis (2003), show that persistent endogenous growth can be driven by the inframarginal rents obtained by competitive firms in the short run. A similar mechanism has been proposed in a series of papers by Boldrin and Levine (see e.g. Boldrin and Levine 2002, 2008), who argue that innovators can profit by selling the first "copy" of their ideas. A common feature of these models is that innovative technological knowledge cannot be immediately used by firms other than the inventor, even in the absence of patent protection. In this sense, these models depart from the traditional assumption that innovative technological knowledge is non rival (Arrow, 1962). The mechanism proposed by Hirshleifer, by contrast, is fully consistent with that assumption. In fact, the swifter and the wider is the adoption of the new technology, the greater are the speculative gains innovators can obtain.

## 2 The baseline model

In this section, we develop a stylised general equilibrium model of endogenous growth where innovations affect the price of a productive asset, creating an opportunity for speculative profits.

There is a unique final good in the economy that can be consumed or used in research. This good is taken as the numeraire. It is produced in a perfectly competitive market, using labour $L$ and an irreproducible asset $T$. The assumption that the asset (e.g., land) is irreproducible allows us to abstract from issues of capital accumulation. The production function is:

$$y(t) = \theta^{k(t)} L^{\alpha} T^{1-\alpha} \qquad \text{with } \theta > 1 \text{ and } 0 < \alpha < 1 \tag{1}$$

where $\theta^{k(t)}$ is total factor productivity at time $t$, the variable $k(t)$ represents the total number of innovations occurred by time $t$, and $\alpha$ is the income share of labour.

The economy is populated by identical, infinitely-lived households whose mass is normalised to one. Households have additive logaritmic intertemporal preferences over consumption flows $c(t)$:[1]

$$u(c) = \int_0^{\infty} \ln\left[c(t)\right] e^{-\rho t} dt, \tag{2}$$

where $\rho$ is the rate of time preference. Each household inelastically supplies one unit of labour and one unit of the irreproducible asset, so $L = 1$ and $T = 1$. Thus, at each point in time output equals total factor productivity: $y(t) = \theta^{k(t)}$.

Technology improves over time as a result of innovative activity. Time is continuous, but we refer to period $k$ as the random time interval between innovation $k$ and innovation $k+1$. We set $k(0) = 0$, normalising total factor productivity at time zero to 1. Each innovation increases total factor productivity by a factor $\theta > 1$, so $\theta$ is the growth factor from one period to the next.

[1] One can easily allow for more general preferences, such as for instance

$$u(c) = \int_0^{\infty} \left[\frac{c(t)^{1-\sigma} - 1}{1-\sigma}\right] e^{-\rho t} dt,$$

where $1/\sigma$ is the intertemporal elasticity of substitution.

In each period $k$, there is a free-entry race for innovation $k+1$. The race starts as soon as innovation $k$ is achieved and disclosed. A number of symmetric risk-neutral firms can participate in this race by investing the final good in independent R&D projects. Each research firm $i$ chooses its R&D effort $n_{i,k}$ to obtain the $k+1$-th innovation. The R&D effort is a flow cost paid until the innovation is achieved. The R&D investment produces an instantaneous probability of success of $\lambda_k n_{i,k}$, where $\lambda_k > 0$ is the productivity of R&D. Since projects are independent, the arrival of innovation $k+1$ follows a Poisson stochastic process with a hazard rate $x_k = \lambda_k n_k$, where $n_k = \sum_i n_{i,k}$ denotes aggregate R&D investment.[2] To guarantee the existence of a steady state, we assume that $\lambda_k = \lambda\theta^{-k}$.[3]

To smooth out the dynamics of aggregate variables and prevent output from jumping up at random time intervals, we posit a continuum of symmetric industries indexed by $\omega \in [0,1]$, as in Grossman and Helpman (1991). Although in each industry $\omega$ the arrival of new innovations is stochastic, by the law of large numbers, total income and consumption are determinate. To guarantee symmetry among industries, we assume that the irreproducible assets are industry specific and that each household supplies an equal share of its total labour supply in each industry. Thus, both labour supply and the supply of the irreproducible asset are fixed at one in each industry. We also assume that the production function (1) and the R&D technology are the same in all industries. As a result, the rate of innovation $x$ will also be the same in all industries. A standard calculation shows that the rate of growth of the economy is $(\theta-1)x$.[4] The focus of the analysis is on a representative industry, so we shall drop the index $\omega$ when there is no risk of confusion.

[2] Firms can adjust their R&D efforts at any point in time, but with a Poisson discovery process they all will choose a constant level of R&D expenditure until someone succeeds, and the next race starts.

[3] In a steady state, the hazard rate $x_k$, and hence the expected duration of time periods, must be constant. Since R&D investment $n_k$ must grow at rate $\theta$ from one period to the next, $\lambda_k$ must decline at rate $\theta$. This explains the knife-edge assumption $\lambda_k = \lambda\theta^{-k}$, which is common to all R&D-driven endogenous growth models (Barro and Sala-i-Martin, 2004, ch. 7).

[4] Aggregate output is $Y(t) = \int_0^1 \theta^{k(\omega,t)} \mathrm{d}\omega$. Since $k(\omega,t)$ jumps up to the next higher integer with

# 3 Speculation and growth

We now determine the equilibrium of the economy under the assumption that markets are perfectly competitive and innovations cannot be appropriated through secrecy or intellectual property rights; that is, all firms active in the product market may freely use the leading technology. Innovators can be rewarded by speculative profits only. These, however, may suffice to sustain steady innovation and growth.

## 3.1 Equilibrium prices

Factor markets are perfectly competitive. Since in period $k$ firms can freely use the leading technology, the wage rate is $w_k = \alpha\theta^k$. The income share of labour is $\alpha$. The remaining share represents the rents accruing to the owners of the irreproducible asset, $R_k = (1-\alpha)\theta^k$.

Let $P_k$ denote the price of the irreproducible asset in period $k$. Investors can perfectly diversify risk by investing in different industries, so in equilibrium the expected return from holding the asset must equal the interest rate $r$. The return to holding the asset is the sum of the rents $R_k$, plus any expected capital gain due to the arrival of innovation $k+1$.

With complete information, when innovation $k+1$ arrives the price would jump to $P_{k+1}$ and stay constant until the next innovation.[5] However, the timing of the innovation is uncertain. The innovator, being "the first in the know", has for a time inside information as to the arrival of the innovation. Thus, he can anticipate the market and obtain speculative profits. This reduces, conversely, the capital gain

a constant instantaneous probability $x$, the rate of growth of the economy is:

$$\frac{\dot{Y}(t)}{Y(t)} = \frac{\int_0^1 \left[\theta^{k(\omega,t)+1} - \theta^{k(\omega,t)}\right] x \mathrm{d}\omega}{\int_0^1 \theta^{k(\omega,t)} d\omega} = (\theta - 1)x.$$

[5] Price increases are permanent when the asset is irreproducible. If the asset was reproducible, by contrast, any increase in its price would stimulate the accumulation of the asset until the price falls back to the asset production cost. The anticipation of this adjustment process would limit changes in the asset price when innovations occur.

obtained by outside investors.

In a noiseless economy, any inside information would be perfectly revealed as soon as its possessor tried to exploit it (Grossman and Stiglitz, 1980). However, it is by now well known that insiders can speculate by hiding behind noise traders (Kyle, 1985). An explicit model of the speculative process should specify the source of noise trading, account for the fact that speculation takes time and the inventor will not disclose the innovation while he is speculating, and determine the optimal timing of disclosure under the threat that the innovation may be duplicated by other firms.[6] Given the state of the art in the literature on insider trading, such a fully microfounded general equilibrium model of speculation seems currently out of reach.

For these reasons, we do not model insider trading explicitly. We assume that the speculative process is instantaneous,[7] and that the innovator captures a share $\gamma > 0$ of the change in the total value of the asset, $P_{k+1} - P_k$, by exploiting his superior information. We take $\gamma$ as a parameter and focus on the macroeconomic consequences of speculation. While the size of $\gamma$ is ultimately an empirical question, our qualitative results hold as long as $\gamma$ is strictly positive and lower than one.

Outside investors anticipate that they will obtain only a fraction $(1 - \gamma)$ of the change in the value of the asset, as the remaining fraction $\gamma$ is reaped by the inventor. This implies that their expected capital gain is $x_k(1-\gamma)(P_{k+1}-P_k)$. The asset pricing equilibrium condition then is:

$$rP_k = R_k + x_k(1 - \gamma)(P_{k+1} - P_k).$$

In a steady state, $x_k$ is constant across periods and the price $P_k$ grows by a factor

[6] This possibility in turn complicates the analysis of the innovation race. As time passes, firms participating in the race must recognise that the probability that somebody else has already innovated and is secretely speculating increases. This affects the reward to successful completion of the R&D project. Hence, the equilibrium investment in R&D will depend on the time passed since the start of the race and on the current value of the asset price, which may signal the occurrence of innovations.

[7] This assumption reflects the fact that speculation is typically much more rapid than innovation. In practice, inside information can hardly be concealed for more than a few days or weeks, whereas achieving an innovation may easily take many years.

$\theta$ from one period to the next. Writing $P_k = \theta^k p$, the asset price equation becomes:

$$rp = (1-\alpha) + x(1-\gamma)(\theta - 1)p. \tag{3}$$

### 3.2 Innovation races

Consider now the equilibrium in innovation races. The prize to the winner of the $k+1$-th race is the speculative profit $\pi_k = \gamma\left(P_{k+1} - P_k\right)$. The expected discounted profit of a generic firm $i$ that invests $n_{i,k}$ units of the final good in that race is

$$\frac{\lambda_k n_{i,k}\pi_k - n_{i,k}}{r + \lambda_k n_k}.$$

Because there is free entry, the zero-profit condition $\lambda_k \pi_k \leq 1$ must hold in equilibrium. This guarantees that the expected return to R&D investment does not exceed the interest rate $r$.

In an equilibrium with positive investment in R&D, the zero-profit condition reduces to $\lambda_k \pi_k = 1$, or, using the fact that $\lambda_k = \lambda\theta^{-k}$ and $P_k = \theta^k p$:[8]

$$\lambda\gamma(\theta - 1)p = 1. \tag{4}$$

### 3.3 Equilibrium growth

From the asset price equation (3) and the zero profit condition (4), one determines an increasing relationship between the interest rate and the rate of innovation:

$$r = \gamma\lambda(\theta - 1)(1-\alpha) + (1-\gamma)(\theta - 1)x. \tag{5}$$

The Euler equation provides another relationship:

$$r = \rho + (\theta - 1)x. \tag{6}$$

These two equations can be solved simultaneously to determine the equilibrium interest rate and the rate of innovation.

[8] With constant returns to research, the equilibrium number of research firms and individual R&D investments are indeterminate, and only aggregate R&D investment is determinate.

**Proposition 1** *In the baseline model, the equilibrium rate of innovation and rate of interest are:*

$$x^* = \lambda(1-\alpha) - \frac{\rho}{\gamma(\theta-1)} \tag{7}$$

$$r^* = (\theta-1)(1-\alpha)\lambda - \left(\frac{1-\gamma}{\gamma}\right)\rho. \tag{8}$$

Proposition 1 implies that for a range of parameter values, speculative profits can sustain steady innovation and growth in a perfectly competitive economy with no monopoly rents. To be precise, the condition is:[9]

$$\gamma > \frac{\rho}{\lambda(1-\alpha)(\theta-1)}. \tag{9}$$

Condition (9) says that the return to R&D when $x = 0$, which is $\gamma(\theta-1)(1-\alpha)/\rho$,[10] exceeds the unit cost of R&D, $1/\lambda$. In equilibrium, the rate of innovation $x$ must then increase up to the point where the returns to R&D become equal the cost.[11]

From (7) and (8) several comparative statics results immediately follow. Quite intuitively, the rate of growth increases with the inventors' ability to appropriate speculative profits, $\gamma$. The rate of growth of the economy depends positively also on the productivity of R&D expenditure, $\lambda$, and the size of innovations, $\theta$, whereas it depends negatively on the rate of time preferences $\rho$. These effects are natural and are the same as in standard Schumpeterian models.

[9] With a constant elasticity instantaneous utility function, the Euler condition would be

$$r = \rho + \sigma(\theta-1)x,$$

where $1/\sigma$ is the intertemporal elasticity of substitution. The analysis does not change if $\sigma > 1-\gamma$. When this inequality is reversed, however, there is no equilibrium with positive growth if condition (9) holds, and there is only an unstable equilibrium if inequality (9) is reversed. In this sense, inequality $\sigma > 1-\gamma$ is another necessary condition, in addition to (9), for growth to be sustainable by speculative profits alone.

[10] When the rate of innovative activity $x$ is zero, the interest rate coincides with the rate of time preference $\rho$. An innovation would then increase the price of the irreproducible asset $(1-\alpha)/\rho$ by an amount equal to $(\theta-1)(1-\alpha)/\rho$, of which the inventor obtains a fraction $\gamma$.

[11] The reason why an increase in $x$ reduces the return to R&D is as follows. There are two opposing effects. First, an increase in $x$ increases expected future rents, raising the current price of the asset. Second, an increase in $x$ makes the interest rate increase in order to satisfy the Euler condition. Since the slope of the zero-profit condition (5) is always lower than that of the Euler condition (6), the latter effect dominates the former.

A novel result is that an increase in the income share of labour reduces growth.[12] This follows from the fact that labour (including human capital) is inalienable. While Eli Whitney and his forethoughted contemporaries could have speculated on the price of slaves, in an economy without slavery only the fraction $(1 - \alpha)$ of the increase in total factor productivity may be reflected in changes in the prices of tradeable assets. The greater is the income share of labour $\alpha$, the more limited are the opportunities for speculative gains.

### 3.4 Social optimum

The social optimum requires both static allocative efficiency and dynamic efficiency. In other words, it requires that an optimal share of income is invested in research. The social trade-off is that an increase in the share of income invested in research makes income grow more quickly, but reduces the share of income consumed.

In the market equilibrium, allocative efficiency is guaranteed at any point in time, as markets are perfectly competitive and there are no static distortions.[13] We now show that in the baseline model speculative profits do not suffice to generate the appropriate incentive to innovate.

**Proposition 2** *In the baseline model, the market equilibrium rate of innovation is always lower than the socially optimal rate of innovation.*

*Proof.* Since there is no capital accumulation, the social problem is stationary. Therefore, the optimal policy must be stationary. With a constant hazard rate $x$, total R&D expenditure is

$$n(t) = \int_0^1 n_k(\omega, t) \mathrm{d}\omega = \frac{xy(t)}{\lambda}.$$

[12] In traditional Schumpeterian models, an increase in the income share of labour reduces the elasticity of demand for innovative goods and hence increases the monopoly price. This allows innovators to obtain higher monopoly rents and so stimulates growth (Barro and Sala-i-Martin, 2004, ch. 7).

[13] The speculative profits earned by inventors are similar to a capital tax, but since the asset is irreproducible, the "tax" does not entail any deadweight loss.

We then have

$$c(t) = \left(1 - \frac{x}{\lambda}\right) y(t) = \left(1 - \frac{x}{\lambda}\right) e^{x(\theta-1)t}.$$

Substituting into the utility function (2) one gets

$$u = \frac{1}{\rho^2} x(\theta - 1) + \frac{1}{\rho} \ln \left(1 - \frac{x}{\lambda}\right)$$

The optimal hazard rate is found by maximising $u$ and is

$$\hat{x} = \lambda - \frac{\rho}{(\theta - 1)}, \tag{10}$$

provided that $(\theta - 1) > \frac{\rho}{\lambda}$. If this inequality is reversed, the optimal policy entails zero R&D investment, so the economy stagnates indefinitely. Comparing $\hat{x}$ with the equilibrium rate of innovation $x^*$, one immediately sees that $x^* < \hat{x}$. ■

This result can be easily understood by contrasting the social and private value of innovations. The social value of innovation $k$ is the discounted increase in total factor productivity, $(\theta^k - \theta^{k-1})/r$. The private value is a fraction $\gamma$ of the increase in the value of the irreproducible asset. There are two reasons why the private value is necessarily lower than the social value. First, inventors obtain only a share $\gamma$ of the increase in the value of the irreproducible asset. Second, the increase in the value of the asset is only a share $(1 - \alpha)$ of the social value of the innovation. The remaining share $\alpha$ increases labour income, which cannot be captured by speculators as labour is not alienable. In section 5 below we show that this underinvestment result holds even in a model that accounts for changes in relative asset prices.

## 4 The weak overinvestment hypothesis

Now we turn to a scenario of proprietary innovations where imitation is prevented by patent protection, so inventors can obtain monopoly rents. We ask whether overinvestment can occur when they obtain speculative profits, too.

In our model, the demand for innovative technology is perfectly rigid, so patent holders need not price discriminate to extract the value of the innovation fully. However, a patent holder's market power may be destroyed by the occurrence of subsequent innovations. To allow for full patent protection (which corresponds to Hirshleifer's case of a "perfectly discriminating patent holder"), we assume that patents can be consolidated into a patent pool. The patent pool licenses the patents to the competitive firms operating in the product market, and each patent holders in the pool obtains his marginal contribution to the pool's total profits.

Let $m$ denote the number of successive patents that can be consolidated in the patent pool. By varying $m$, one can then capture different degrees of patent protection. The case where $m$ is arbitrarily large corresponds to full protection, and the case $m = 1$ to the more plausible assumption that successive patent holders compete with each other in the market for technology.[14] When $m = 0$, we are back to the model with no patent protection.

Patent holders compete *à la* Bertrand. In period $k$, innovator $k - m$ has just been excluded from the patent pool and so must stand ready to license his technology at a zero royalty rate. The patent pool, which licenses the most productive technology, will then charge an aggregate royalty rate per unit of output equal to:[15]

$$\varphi_k = 1 - \frac{1}{\theta^m} \tag{11}$$

and obtain an aggregate profit of $\theta^k - \theta^{k-m}$. With full patent protection, the royalty rate is $1 - \frac{1}{\theta^k}$, as the initial technology, of vintage 0, is in the public domain.

Each past innovator $j$ who still participates in the patent pool obtains a share of this aggregate profit equal to his marginal contribution, $\left(\theta^j - \theta^{j-1}\right) / \left(\theta^k - \theta^{k-m}\right)$. When a new innovation arrives, each past innovator's share in the patent pool's profit decreases, but total profit increases in such a way that individual profit stays

[14] A standard argument, based on the Arrow replacement effect, implies that the incumbents do not conduct any research and hence are systematically replaced by an outsider.

[15] To understand this formula, notice that the net output of a perfectly competitive firm that licenses the state-of-the-art technology is

$$\theta^k L^\alpha T^{1-\alpha} - \left(1 - \frac{1}{\theta^m}\right) \theta^k L^\alpha T^{1-\alpha} = \theta^{k-m} L^\alpha T^{1-\alpha},$$

i.e., the same as if the firm used the technology of vintage $k - m$, which is less productive but does not command any royalty.

constant. However, after $m$ successive innovations the patent holder is excluded from the patent pool and his profits vanish. Thus, innovator $j$ obtains monopoly rents equal to $\left(\theta^j - \theta^{j-1}\right)$ for $m$ periods.

In addition, he can also obtain speculative profits. With patent protection, the wage rate and the rents obtained by the irreproducible asset become $w_k = \alpha\theta^{k-m}$ and $R_k = (1-\alpha)\theta^{k-m}$, respectively. Thus, each factor is rewarded as if the technology of vintage $k-m$, instead of the state-of-the-art technology of vintage $k$, were used. Relative to the baseline model, rents, and hence the equilibrium price of the asset, are scaled down by a factor $\theta^{-m}$. When an innovation occurs, the asset price jumps up by a factor $\theta$ creating again an opportunity to speculate, but speculative gains are also scaled down by a factor $\theta^{-m}$. Thus, patent protection provides monopoly rents, but crowds out speculative profits.

The total discounted profits accruing to innovator $k$ are therefore

$$\pi_k = \left(\theta^k - \theta^{k-1}\right) \frac{1 - \left(\frac{x}{r+x}\right)^m}{r} + \gamma\left(\theta^k - \theta^{k-1}\right)\theta^{-m} p, \tag{12}$$

where the first terms captures monopoly rents and the second speculative profits. Equation (12) shows that an increase in $m$ increases monopoly rents but decreases speculative gains. With full patent protection ($m = \infty$), monopoly rents are maximised: the $k$-th innovator obtains a permanent flow of profits of $\left(\theta^k - \theta^{k-1}\right)$, which equals the full social value of his innovation. However, such a fully protected patent holder cannot obtain speculative profits at all. With full patent protection each factor is rewarded as if the technology of vintage 0 were used, rather than the state-of-the-art technology. All the productivity increase is reaped by the patent pool. As a result, asset prices do not change at all when a new innovation arrives. Speculative profits are crowded out entirely. The innovator obtains exactly the full social value of his innovation, so the equilibrium rate of innovation is just socially optimal, not higher. This means that the weak overinvestment hypothesis does not hold with full patent protection.

However, the total incentive to innovate may decrease with the degree of patent protection $m$. As we have seen above, reducing $m$ decreases the magnitude of monopoly rents, but increases the size of speculative profits. For any value of $\gamma > 0$, one can find a value of $x$ small enough that the latter effect dominates the former, implying that the total incentive to innovate $\pi_k$ is largest when $m = 1$. The intuition is that an increase in $m$ prolongs the expected duration of the period over which innovator $k$, not having been displaced by subsequent innovations yet, can collect monopoly rents. When $x$ is close to zero, however, that period is very long even if patents are not consolidated, i.e. if $m = 1$. In this case, allowing for patent consolidation has a second order effect on monopoly rents, but a negative first order effect on speculative gains. This opens the possibility of overinvestment in R&D when patent protection is limited.

**Proposition 3** *With partial patent protection, the market equilibrium rate of innovation can exceed the socially optimal rate of innovation.*

*Proof.* To fix ideas, consider the case $m = 1$. The free entry condition becomes:

$$\frac{(\theta - 1)}{r + x} + \gamma \left(1 - \frac{1}{\theta}\right) p = \frac{1}{\lambda}.$$

To show the possibility of overinvestment in R&D, we contrast the conditions for the market equilibrium and the socially optimal rate of innovation to be positive. The socially optimal rate of innovation (10) is positive only if

$$(\theta - 1) > \frac{\rho}{\lambda}$$

The equilibrium rate of innovation, by contrast, is positive only if

$$(\theta - 1) \left[1 + \gamma \frac{(1 - \alpha)}{\theta}\right] > \frac{\rho}{\lambda}.$$

Clearly, as long as $\gamma > 0$ the market equilibrium rate of innovation can be positive for parameter values for which the socially optimal rate is zero. This suffices to prove the possibility of overinvestment. ■

## 5 The strong overinvestment hypothesis

Now we return to the assumption that there is no patent protection. In the baseline model, speculative profits by themselves cannot generate an excessive incentive to innovate. However, the baseline model might underestimate the potential for speculative gains. In particular, it does not capture the redistributive effects of technical change discussed by Hirshleifer in his "locomotive technology" example. A feature of that example is that there are various assets in the economy and innovations are asset specific, meaning that they increase the productivity of some assets but not that of others. Thus, the occurrence of the innovation appreciates certain assets but depreciates others, amplifying the opportunities for speculation.

In this section, we modify the baseline model so as to capture these effects and allow for changes in *relative* asset prices. Nevertheless, we show that there can never be overinvestment in R&D when inventors are rewarded by speculative profits only.

### 5.1 Asset specific innovations

Assume that each industry $\omega$ comprises two sectors, indexed by $v$. In each sector, the final good is produced using labour and a sector-specific, irreproducible asset. We normalise the supply of both assets to one and denote their prices by $P_{v,k}$. The supply of labour is fixed and equal to one in each industry. However, labour can now freely move across the two sectors of an industry.

In sector $v$, the production function is given by:

$$y_v(t) = \theta^{h_v(t)} L_v^{\alpha}(t) \tag{13}$$

where $L_v$ is labour input and $h_v(t)$ is a technological index that depends on the number of past innovations. We now specify how this technological index is determined.

In each industry, the technological frontier corresponds to a total factor productivity equal to $\theta^{k(t)}$, where $k(t)$ is the total number of past innovations occurred in the industry by time $t$. As in the baseline model, the variable $k(t)$ represents the industry-wide stock of knowledge, which all subsequent innovations build on in a cumulative way. However, each innovation is now targeted to a specific asset, and hence to a specific sector of the industry. That is, innovation $k+1$ raises total factor productivity to $\theta^{k+1}$ only in the sector in which it occurs, leaving total productivity unchanged in the other sector.

With these assumptions, the two sectors never share the same technology. In the advanced sector, i.e. the sector where the latest innovation has occurred, we have $h(t) = k(t)$. In the less advanced sector, by contrast, $h(t)$ equals the latest period in which an innovation occurred there. The technological gap between the two sectors depends on whether sectors alternate in leading, or several innovations occur in a row in the same sector. This is determined endogenously in equilibrium, as we shall see below.

The assumption that innovations are asset specific is crucial to generate changes in relative asset prices. Since factor productivity increases only in the sector where the innovation has occurred, labour flows from the less productive sector to the more productive one. The rents in the advanced sector increase because of the increase in productivity, and because of the inflow of labour. The rents in the less advanced sector, by contrast, decrease because of the outflow of labour. These changes in relative asset prices amplify the opportunities for speculative profits as compared to the baseline model.

Like in the baseline model, we assume that the innovator can use his inside information about the arrival of the innovation to capture a share $\gamma$ of the increase in the value of the irreproducible asset in the sector where the innovation occurred. For simplicity, we rule out short sales.[16]

Let $n_{v,k} = \sum_i n_{i,v,k}$ denote aggregate R&D investment per unit of time in period $k$ targeted to sector $v$. Then, the $k+1$-th innovation occurs in sector $s$ according

[16] However, it can be shown that the main result of this section, i.e. Proposition 4, continues to hold even allowing for short sales.

to a Poisson process with a hazard rate $x_{v,k} = \lambda_k n_{v,k}$. We continue to assume that $\lambda_k = \lambda\theta^{-k}$.

### 5.2 Equilibrium

Research firms now choose both the level of the R&D investment and the sector they target. While we cannot rule out the possibility of multiple equilibria, we focus on the case where the opportunity for speculative profits are largest. Evidently, this requires that all research is directed to the less advanced sector, so that in equilibrium the sectors systematically alternate in leading.[17] The following lemma guarantees the existence of such an equilibrium:

**Lemma 1** *There exists an equilibrium in which $x_{v,k} = 0$ whenever $h_v = k$.*

*Proof.* See appendix A.

Intuitively, the change in asset prices is largest if the innovation occurs in the less advanced sector, where innovation $k+1$ would raise total factor productivity by two steps rather than one. With constant returns to scale in research, profit-maximising research firms will target the sector where the arrival of the innovation generates the greatest change in asset prices. Since sectors alternate in leading, we shall denote by 1 the more advanced sector and by -1 the less advanced one.

Factor markets are perfectly competitive, so the wage rate equals the marginal productivity of labour:

$$w_k = \alpha\theta^k L_1^{\alpha-1} = \alpha\theta^{(k-1)} L_{-1}^{\alpha-1}. \tag{14}$$

The marginal productivity of labour is equalised across sectors at each point in time, and the allocation of labour is efficient. Together with the labour market clearing condition $L_1 + L_{-1} = 1$, equation (14) can be solved to yield:

$$L_1 = \frac{\eta}{1+\eta} \qquad \text{and} \qquad L_{-1} = \frac{1}{1+\eta}$$

[17] If in equilibrium all research was directed to the leading sector, one would effectively be back to the baseline model.

where $\eta \equiv \theta^{\frac{1}{1-\alpha}} > 1$. Clearly, $L_1 > L_{-1}$. Since sectors alternate in leading, when a new innovation occurs labour instantaneously flows to the sector where productivity has increased.

The rents that accrue to the owners of the sector-specific assets are:

$$\begin{aligned} R_{1,k} &= (1-\alpha)\left(\frac{\eta}{1+\eta}\right)^{\alpha}\theta^{k} \\ R_{-1,k} &= \frac{(1-\alpha)}{\theta}\left(\frac{1}{1+\eta}\right)^{\alpha}\theta^{k}. \end{aligned} \tag{15}$$

It can be shown that the rate of growth of the economy is still $(\theta-1)x$.[18]

Like in the baseline model, the return to asset $v$ is the sum of the rents earned by the asset plus any expected capital gain or loss. In equilibrium, the rate of return must equal the interest rate $r$, implying:

$$rP_{1,k} = R_{1,k} + x_k(P_{-1,k+1} - P_{1,k}),$$

[18] Substituting the equilibrium labour inputs into the production function, one gets the equilibrium outputs:

$$y_{1,k} = \left(\frac{\eta}{1+\eta}\right)^{\alpha}\theta^{k} \qquad \text{and} \qquad y_{-1,k} = \left(\frac{1}{1+\eta}\right)^{\alpha}\theta^{k-1}.$$

In each industry, total output

$$\begin{aligned} y_k &= y_{1,k} + y_{-1,k} \\ &= \left(\frac{1}{1+\eta}\right)^{\alpha}\left(\eta^{\alpha} + \frac{1}{\theta}\right)\theta^{k} \end{aligned}$$

grows at rate $\theta - 1$ from one period to the next. Since there is a continuum of industries, however, aggregate variables grow smoothly. Summing across industries, aggregate output is

$$\begin{aligned} Y(t) &= \int_0^1 y_k(\omega,t)\mathrm{d}\omega \\ &= \left(\frac{1}{1+\eta}\right)^{\alpha}\left(\eta^{\alpha} + \frac{1}{\theta}\right)G(t), \end{aligned}$$

where $G(t) \equiv \int_0^1 \theta^{k(\omega,t)}d\omega$ is an average productivity index that increases over time with technical progress. The rate of growth of output is the rate of growth of the average productivity index, $G(t)$. To calculate it, notice that $k(\omega,t)$ jumps up to the next higher integer with a constant instantaneous probability $x$. Hence:

$$\begin{aligned} \dot{G}(t) &= \int_0^1 \left[\theta^{k(\omega,t)+1} - \theta^{k(\omega,t)}\right] x\mathrm{d}\omega \\ &= (\theta-1)xG. \end{aligned}$$

and

$$rP_{-1,k} = R_{-1,k} + x_k(1-\gamma)(P_{1,k+1} - P_{-1,k}),$$

since only a fraction $(1-\gamma)$ of the capital gain $(P_{1,k+1} - P_{-1,k})$ accrues to outside investors; the remaining fraction $\gamma$ is the reward to the innovator.

In a steady state, the asset price equations become:

$$\begin{aligned} rp_1 &= R_1 + x(\theta p_{-1} - p_1) \\ rp_{-1} &= R_{-1} + x(1-\gamma)(\theta p_1 - p_{-1}), \end{aligned} \quad (16)$$

where $P_{s,k} \equiv \theta^k p_s$ and $R_{s,k} \equiv R_s \theta^k$. These equations can be solved to express $p_1$ and $p_{-1}$ as a function of $x$.

For future reference, we note that the occurrence of the innovation increases the price of the asset used in the sector where productivity increases, but decreases that of the other asset.

**Lemma 2** *$P_{1,k+1} > P_{-1,k}$ and $P_{-1,k+1} < P_{1,k}$.*

*Proof.* See appendix A.

The speculative gains accruing to the $k+1$-th inventor are $\pi_k = \gamma(P_{1,k+1} - P_{-1,k})$, or $\pi_k = \gamma(\theta p_1 - p_{-1})\theta^k$. The zero-profit condition in innovation races then becomes

$$\gamma\lambda(\theta p_1 - p_{-1}) = 1. \quad (17)$$

Since $p_1$ and $p_{-1}$ are a function of $x$, the zero-profit condition determines a relationship between the interest rate and the rate of innovation. Like in the baseline model, the Euler equation provides another relationship, which together with the zero-profit condition uniquely determines the equilibrium interest rate and the rate of innovation.

Like in the baseline model, speculative profits can sustain innovation and growth. The necessary and sufficient condition is that the returns to R&D when no further innovation is anticipated exceed the unit cost of R&D. When $x = 0$, asset prices reduce to $p_1 = R_1/\rho$ and $p_{-1} = R_{-1}/\rho$. Thus, growth can be sustained by speculative profits if and only if

$$\gamma \frac{(\theta R_1 - R_{-1})}{\rho} > \frac{1}{\lambda}. \tag{18}$$

It is immediate to verify that condition (18) is weaker than condition (9), confirming that the potential for speculation is higher in the two-asset model.

### 5.3 Comparison with social optimum

Now we are in a position to compare once again the market equilibrium with the social optimum. Although the potential for speculation is higher than in the baseline model, we have:

**Proposition 4** *In the model with asset specific innovations, the market equilibrium rate of innovation is still always lower than the socially optimal rate.*

*Proof.* It can be easily confirmed that the equilibrium rate of innovation is largest when $\gamma = 1$. In this case, the innovator captures all the increase in the value of the asset that appreciates when the innovation arrives. Asset equilibrium prices then become

$$\begin{aligned} p_1 &= \frac{rR_1 + x\theta R_{-1}}{r(r+x)} \\ p_{-1} &= \frac{R_{-1}}{r}. \end{aligned}$$

Notice that the price of the less productive asset is independent of $x$ and always equals the discounted value of the rents $R_{-1}$, as all future capital gains are appropriated by the innovator. The price of the more productive asset, by contrast, decreases with $x$. This follows from the fact that $R_1 > \theta R_{-1}$. The intuitive reason is that holders of the more productive asset suffer a capital loss when the new innovation arrives in the other sector, causing a reallocation of labour across sectors. It follows that the incentive to innovate, $\theta p_1 - p_{-1}$, is now a decreasing function of $x$. Unlike the baseline model, the zero profit condition now determines a decreasing relationship between the interest rate $r$ and the rate of innovation $x$:

$$x = r\frac{\lambda(\theta R_1 - R_{-1}) - r}{r - \lambda(\theta^2 - 1)R_{-1}}. \tag{19}$$

Together with the Euler equation (6), this condition determines the equilibrium interest rate $r^*$ and the rate of innovation $x^*$. Clearly, an upper bound on the equilibrium rate of innovation is placed by inequality $r < \lambda(\theta R_1 - R_{-1})$, which is a necessary condition for $x^* > 0$. (Notice that condition (18) implies that the denominator of (19) must be positive.) Together with the Euler equation, this implies:

$$x^* < \lambda\frac{\theta R_1 - R_{-1}}{(\theta - 1)} - \frac{\rho}{(\theta - 1)}. \tag{20}$$

The socially optimal rate of innovation can be calculated proceeding as in the baseline model. Optimality requires that the static allocative efficiency condition (14) holds. Clearly, the social planner will direct all the research to the less advanced sector, where there is more to gain from innovating. Thus, along the optimal path sectors will alternate in leading, as in the market equilibrium we have been focusing on. The optimal resolution to the dynamic trade-off between current and future consumption lead to the following optimal rate of innovation:[19]

$$\hat{x} = \lambda\frac{R_1 + R_{-1}}{(1 - \alpha)} - \frac{\rho}{(\theta - 1)}. \tag{21}$$

We now prove that $x^* < \hat{x}$. From (20) and (21), it follows that a sufficient condition for $x^* < \hat{x}$ is:

$$\Delta(\alpha, \theta) \equiv \theta R_1 - R_{-1} - \frac{R_1 + R_{-1}}{1 - \alpha}(\theta - 1) \leq 0.$$

Substituting for $R_1$ and $R_{-1}$, $\Delta(\alpha, \theta)$ can be written as:

$$\Delta(\alpha, \theta) = -\left(\frac{\theta^{\frac{1}{1-\alpha}}}{1 + \theta^{\frac{1}{1-\alpha}}}\right)^{\alpha} \theta^{-\frac{1}{1-\alpha}} \left[(\theta\alpha - 1)\,\theta^{\frac{1}{1-\alpha}} - (\alpha - \theta)\right].$$

[19] For simplicity, the calculation is based on the assumption that the initial conditions conform to the steady state properties. That is, we have assumed that initially in each industry one sector has a one-step technological lead over the other, as is always true in the steady state.

Thus, the sufficient condition becomes:

$$H(\alpha,\theta) = (\theta\alpha - 1)\,\theta^{\frac{1}{1-\alpha}} - (\alpha - \theta) \geq 0$$

Since $H(\alpha, 1) = 0$, the sufficient condition becomes $H'_\theta(\alpha,\theta) \geq 0$. We calculate:

$$H'_\theta(\alpha,\theta) = \frac{\alpha\,(2-\alpha)\,\theta^{\frac{1}{1-\alpha}} - \theta^{\frac{\alpha}{1-\alpha}} + (1-\alpha)}{1-\alpha}.$$

This implies that a sufficient condition for $x^* < \hat{x}$ is that $K(\alpha,\theta) \equiv \left[\alpha\,(2-\alpha)\,\theta^{\frac{1}{1-\alpha}} - \theta^{\frac{\alpha}{1-\alpha}} + (1-\alpha)\right] \geq 0$. We have $K(\alpha, 1) = \alpha\,(1-\alpha) \geq 0$. Thus, the sufficient condition can be restated as $K'_\theta(\alpha,\theta) \geq 0$. Finally, we verify:

$$K'_\theta(\alpha,\theta) = \frac{\alpha\theta^{\frac{\alpha}{1-\alpha}}}{(1-\alpha)}\left[(2-\alpha) - \theta^{-1}\right] \geq 0$$

since $2 - \alpha \geq 1$ and $\theta^{-1} < 1$.■

The intuition is as follows. The reason why there is more scope for speculation in the model with asset specific innovations is that the reallocation of labour across sectors amplifies changes in asset prices. Clearly, the effect of labour reallocation can be strong only if the income share of labour is large. However, when $\alpha$ is large speculative profits must be small as compared to the social value of innovations. Thus, the fact that labour is a non alienable asset still produces an under-investment result.

# 6 Conclusion

The main contribution of the paper is to formalise Hirshleifer's insight that innovative activity can be stimulated by speculative profits. These can be obtained by inventors even in the absence of patent protection thanks to their inside information as to the occurrence of the innovation. We have developed a model of a perfectly competitive economy in which steady innovation and growth can be sustained by such speculative gains. While unexpected radical innovations may evidently have a large impact on asset prices, our model shows that even the steady, predictable flow of innovations postulated in models of equilibrium growth can create sufficient speculative opportunities.

The analysis has revealed a trade-off between monopoly rents and speculative gains. Strengthening patent protection increases the former but decreases the latter. As a result, the total incentive to innovate may well be highest for intermediate degrees of patent protection. Indeed, we have shown that in our model overinvestment is possible when patent protection is somewhat limited, but there can never be overinvestment in R&D with either zero or maximum patent protection. This confirms Hirshleifer's weak overinvestment hypothesis, although the mechanism is subtler than he conjectured.

The strong overinvestment hypothesis, on the contrary, is not confirmed by our analysis. Inventors rewarded by speculative gains only are necessarily undercompensated. The reason is that the social value of innovations is the total increase in factor productivity, which includes the increase in the productivity of labour. Since labour is non alienable, and hence cannot be the object of speculation, inventors can at most obtain only a share of what they contributed to society.

In our model, speculative profits by themselves could generate excessive incentives to invest in R&D only in a slave economy with asset specific innovations.[20] In a slave economy, labour is alienable and so speculators can in principle capture the entire social value of innovations. In addition, with asset specific innovations, the opportunities for speculation are enhanced by the change in relative asset prices that innovations bring about. Both assumptions (i.e., slavery and asset specific innovations) are necessary to generate overinvestment when inventor are remunerated by speculative profits only.

[20] See appendix B.

## References


[1] Arrow, K., 1962, Economic welfare and the allocation of resources to invention," in R. Nelson (ed.), The rate and direction of economic activity, Princeton University Press, N.Y.

[2] Barro, R. and X. Sala-i-Martin, 2004, *Economic growth*, 2nd edition, Cambridge, MA: MIT Press.

[3] Bester, H. and E. Petrakis, 2003, Wages and productivity growth in a competitive industry, *Journal of Economic Theory*, 109, 52-69.

[4] Boldrin, M. and D. Levine, 2002, The case against intellectual monopoly, *American Economic Review Papers and Proceedings*, 92, 209-212.

[5] Boldrin, M. and D. Levine, 2008, Perfectly competitive innovations, *Journal of Monetary Economics*, 55, 435-453.

[6] Grossman, G. M. and E. Helpman, 1991, Quality ladders in the theory of growth, *Review of Economic Studies*, 58, 43-61.

[7] Grossman, S. and J. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review*, 70, 393-408.

[8] Hellwig, M. and A. Irmen, 2001, Endogenous technical change in a competitive economy, *Journal of Economic Theory*, 101, 1-39.

[9] Hirshleifer, J., 1971, The private and social value of information and the reward to inventive activity, *American Economic Review*, 61, 561-574.

[10] Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica*, 53, 1315-1335.

## APPENDIX A: Proofs of the lemmas

**Proof of Lemma 1**. We must show that there exists an equilibrium where all research is targeted to the less advanced sector. If research were directed to the more advanced sector, speculative profits would be a fraction $\gamma$ of $(P_{2,k+1} - P_{1,k}) = (\theta p_2 - p_1)\theta^k$, where $p_2$ is the asset price in a sector that leads by two steps. If research is directed to the less advanced sector, by contrast, speculative profits are a fraction $\gamma$ of $(P_{1,k+1} - P_{-1,k}) = (\theta p_1 - p_{-1})\theta^k$. It follows that the incentive to invest is the less advanced sector is greater than in the advanced sector if

$$\theta p_2 - p_1 < \theta p_1 - p_{-1}. \tag{A.1}$$

To show that this condition is satisfied, we must determine $p_2$. This, however, may in turn depend on the asset price when a sector is leading by three, four or more steps. In general, in a steady state asset prices are determined by the following arbitrage conditions

$$\begin{aligned}
rp_1 &= R_1 + x_1(1-\gamma)(\theta p_2 - p_1) + x_{-1}(\theta p_{-1} - p_1) \\
rp_2 &= R_2 + x_2(1-\gamma)(\theta p_3 - p_2) + x_{-2}(\theta p_{-1} - p_2) \\
&\ \ ... \\
rp_i &= R_i + x_i(1-\gamma)(\theta p_{i+1} - p_i) + x_{-i}(\theta p_{-1} - p_i) \\
&\ \ ... \\
rp_{-1} &= R_{-1} + x_{-1}(1-\gamma)(\theta p_1 - p_{-1}) + x_1(\theta p_{-2} - p_{-1}) \\
rp_{-2} &= R_{-2} + x_{-2}(1-\gamma)(\theta p_1 - p_{-2}) + x_2(\theta p_{-3} - p_{-2}) \\
&\ \ ... \\
rp_{-i} &= R_{-i} + x_{-i}(1-\gamma)(\theta p_1 - p_{-i}) + x_i(\theta p_{-i-1} - p_{-i}) \\
&\ \ ...
\end{aligned}$$

where the rents $R_i$ are determined by the condition that the marginal productivity of labour must be equalised across sectors. This implies

$$w_{i,k} = \alpha\theta^k L_i^{\alpha-1} = \alpha\theta^{(k-i)} L_{-i}^{\alpha-1}.$$

Together with the labour market clearing condition $L_i + L_{-i} = 1$, this condition yields:

$$L_i = \frac{\eta^i}{1+\eta^i} \qquad \text{and} \qquad L_{-i} = \frac{1}{1+\eta^i}$$

The rents then become:

$$\begin{aligned} R_i &= (1-\alpha)\left(\frac{\eta^i}{1+\eta^i}\right)^\alpha \\ R_{-i} &= \frac{(1-\alpha)}{\theta}\left(\frac{1}{1+\eta^i}\right)^\alpha. \end{aligned}$$

To confirm that there is an equilibrium with $x_i = 0$, we must consider out-of-equilibrium beliefs. Assuming rational expectations, we consider a candidate equilibrium where $x_{-1}(= x) > 0$ and $x_{-2} > 0$. (Values of $x_{-i}$ for $i > 2$ are irrelevant.) In this candidate equilibrium, the arbitrage conditions become:

$$\begin{aligned} rp_1 &= R_1 + x_{-1}(\theta p_{-1} - p_1) \\ rp_2 &= R_2 + x_{-2}(\theta p_{-1} - p_2) \\ rp_{-1} &= R_{-1} + x_{-1}(1-\gamma)(\theta p_1 - p_{-1}) \\ rp_{-2} &= R_{-2} + x_{-2}(1-\gamma)(\theta p_1 - p_{-2}). \end{aligned} \tag{A.2}$$

Since $x_{-1} > 0$ and $x_{-2} > 0$, the corresponding zero-profit conditions must hold as equalities. Thus, we have

$$\begin{aligned} \gamma\left(\theta p_1 - p_{-1}\right) &= \frac{1}{\lambda} \\ \gamma(\theta p_1 - p_{-2}) &= \frac{1}{\lambda}, \end{aligned}$$

which implies

$$p_{-1} = p_{-2}.$$

Because $R_{-2} < R_{-1}$, for $p_{-1}$ to equal $p_{-2}$ it must be

$$x_{-2} > x_{-1}.$$

To proceed, notice that the system (A.2) is recursive, as the asset price conditions relative to $p_1$ and $p_{-1}$ are independent of the others. Notice also that $p_2$ is a decreasing function of $x_{-2}$. Since in the candidate equilibrium $x_{-2}$ cannot be lower than $x_{-1}$, a sufficient condition for inequality (A.1) to hold is that the inequality be satisfied when $x_{-2} = x_{-1}$.

Assuming that $x_{-2} = x_{-1} = x$, asset prices are

$$\begin{aligned} p_1 &= \frac{R_1 + x\theta p_{-1}}{r + x} \\ p_2 &= \frac{R_2 + x\theta p_{-1}}{r + x}. \end{aligned}$$

This can be rewritten as

$$\begin{aligned} p_1 &= \frac{r\bar{p}_1 + x\tilde{p}_1}{r + x} \\ p_2 &= \frac{r\bar{p}_2 + x\tilde{p}_2}{r + x}, \end{aligned}$$

where $\bar{p}_i = R_i/r$ and $\tilde{p}_i = \theta p_{-1}$. That is, $p_i$ is a weighted average of $\bar{p}_i$ and $\tilde{p}_i$, with weights equal to $r$ and $x$, respectively. Simple algebra shows that inequality (A.1) is satisfied both when $p_i = \bar{p}_i$ and $p_i = \tilde{p}_i$. Since $p_i$ is a weighted average of $\bar{p}_i$ and $\tilde{p}_i$, inequality (A.1) must always hold. This completes the proof of lemma 1.

**Proof of Lemma 2**. The first part of the lemma is obvious, so it suffices to show that $p_1 > \theta p_{-1}$. The system (16) is linear in $p_1$ and $p_{-1}$ and the matrix of coefficients has full rank (assuming that the transversality condition $r > (\theta - 1)x$ holds, which is always true in equilibrium). Thus, the system implicitly defines $p_1$ and $p_{-1}$ as continuous and differentiable functions of $x$ and $r$:

$$\begin{aligned} p_1 &= \frac{R_1 \left[r + x(1 + \theta - \gamma)\right] - (R_1 - R_{-1})\theta x}{\left[r^2 + (2 - \gamma)rx - (\theta^2 - 1)(1 - \gamma)x^2\right]} \\ p_{-1} &= \frac{R_1 \left[r + x(1 + \theta - \theta\gamma)\right] - (R_1 - R_{-1})(r + x)}{\left[r^2 + (2 - \gamma)rx - (\theta^2 - 1)(1 - \gamma)x^2\right]}. \end{aligned} \tag{22}$$

Notice also that $p_{-1}$ is always increasing in $x$. On the other hand, it is clear from (14) that $p_1$ would be non-increasing in $x$ if $p_1 \leq \theta p_{-1}$.

Now suppose to the contrary that inequality $p_1 > \theta p_{-1}$ does not hold for some $x_0$. Then, the inequality would not hold for any $x \geq x_0$. However, consider the case where $x$ is largest, i.e. the case $\gamma = 1$. In this case, we have

$$\begin{aligned} p_1 &= \frac{rR_1 + x\theta R_{-1}}{r(r+x)} \\ p_{-1} &= \frac{R_{-1}}{r}, \end{aligned}$$

so

$$\begin{aligned} p_1 - \theta p_{-1} &= \frac{rR_1 + x\theta R_{-1} - \theta(r+x)R_{-1}}{r(r+x)} \\ &= \frac{R_1 - \theta R_{-1}}{(r+x)} > 0, \end{aligned}$$

where the inequality holds as $R_1 > \theta R_{-1}$. This contradiction completes the proof of the lemma.

## APPENDIX B: A slave economy

To verify that the cause of underinvestment with no patent protection is, indeed, that labour is non alienable, in this appendix we consider the (fortunately) hypothetical case of a slave economy.

Assume that all labour is slave labour. Each household owns one slave and one unit of each non labour asset. For simplicity, we assume that the inventor captures the same fraction $\gamma$ of the increase in the value of both slaves and non-labour assets.

The value of a slave, $S_k$, is determined by an asset price condition similar to (3):

$$rS_k = w_k + x(1-\gamma)(S_{k+1} - S_k).$$

In a steady state where $S_k \equiv s\theta^k$, this condition reduces to:

$$rs = \alpha + x(1-\gamma)(\theta - 1)s.$$

The static equilibrium conditions are the same as in the case where labour is non alienable, and the price of non-labour assets are still determined by the system (16). Since $w_k = \left(\frac{\eta}{1+\eta}\right)^{\alpha-1} \alpha\theta^k$, the value of a slave is given by

$$rs = \left(\frac{\eta}{1+\eta}\right)^{\alpha-1} \alpha + x(1-\gamma)(\theta-1)s.$$

Total speculative gains now are $\gamma\left[(P_{1,k+1} - P_{-1,k}) + (S_{k+1} - S_k)\right]$.

Now we have:

**Proposition 5** *In a slave economy with asset specific innovations, the market equilibrium rate of innovation can exceed the socially optimal rate of innovation.*[21]

*Proof.* Let us focus once again on the limiting case $\gamma = 1$. In the market equilibrium, the free entry condition is

$$\frac{r\theta R_1 + \left[x(\theta^2 - 1) - r\right] R_{-1}}{r(r+x)} + (\theta - 1)\frac{\left(\frac{\eta}{1+\eta}\right)^{\alpha-1} \alpha}{r} = \frac{1}{\lambda}.$$

[21] It is easy to show that allowing for slavery in the baseline model would not suffice to vindicate the strong overinvestment hypothesis. Both slavery and asset specific innovations are necessary to generate overinvestment when inventors are remunerated by speculative profits only.

It follows that the market equilibrium rate of growth is positive if

$$\theta R_1 - R_{-1} + (\theta - 1)\left(\frac{\eta}{1+\eta}\right)^{\alpha-1}\alpha \geq \frac{\rho}{\lambda}.$$

On the other hand, the optimal growth rate (20) is positive only if

$$\frac{(\theta - 1)(R_1 + R_{-1})}{(1-\alpha)} \geq \frac{\rho}{\lambda}.$$

Denote by

$$F(\alpha, \theta) \equiv \theta R_1 - R_{-1} - (\theta - 1)\frac{R_1 + R_{-1}}{1-\alpha} + (\theta - 1)\alpha\left(\frac{\theta^{\frac{1}{1-\alpha}}}{1+\theta^{\frac{1}{1-\alpha}}}\right)^{\alpha-1}$$

the difference between the left-hand sides of the two inequalities above. We must show that $F(\alpha, \theta) > 0$ for at least some values of $\theta$ and $\alpha$. In fact, simple algebra shows that

$$F(\alpha, \theta) = (1-\alpha)\left(\frac{\theta^{\frac{1}{1-\alpha}}}{1+\theta^{\frac{1}{1-\alpha}}}\right)^{\alpha}\left(1 - \theta^{-\frac{\alpha}{1-\alpha}}\right) > 0,$$

for $1 > \alpha > 0$. This shows that the market equilibrium rate of innovation can be positive for parameter values for which the socially optimal rate is zero. This suffices to demonstrate the possibility of overinvestment. ■

The intuitive explanation is that the social value of innovation $k+1$ equals the total increase in the value of labour and non-labour assets that it brings about, which is $(\theta - 1)(P_{1,k} + P_{-1,k} + S_k)$. The private value of innovation $k+1$, by contrast, is $\gamma\left[\theta P_{1,k} - P_{-1,k} + (\theta - 1)S_k\right]$. When $\gamma = 1$, the private value exceeds the social value by an amount equal to $(P_{1,k} - \theta P_{-1,k})$, which is the asset depreciation in the formerly advanced sector. This is positive by Lemma 2.