What Causes Over-investment in R&D in Endogenous Growth Models?*

Vincenzo Denicolò
University of Leicester, University of Bologna, and CEPR

Piercarlo Zanchettin
University of Leicester

January 12, 2014

Abstract

Endogenous growth models may exhibit either under- or over-investment in R&D. The possibility of over-investment is generally attributed to a business-stealing effect that arises as the latest innovator destroys and/or appropriates previous incumbent's rents. We argue that this conventional wisdom is misleading. In standard models, business stealing by itself cannot result in excessive R&D. We explain the other effects that must be at work here, thus contributing towards a better understanding of when and why the market may be biased towards excessive R&D.

Keywords: Endogenous growth; over-investment in R&D; business-stealing effect; monopoly distortion effect; R&D congestion effect

JEL classification number: O10, O30

*E-mail addresses: vincenzo.denicolo@unibo.it (Vincenzo Denicolò), pz11@le.ac.uk (Piercarlo Zanchettin). We are grateful to the Editor, Martin Cripps, three anonymous referees, Gene Grossman, Antonio Minniti and Ludovic Renou for useful comments and discussions. The usual disclaimer applies.

It is a well known fact that R&D-driven endogenous growth models may exhibit either under- or over-investment in R&D. The possible origins of under-investment in R&D are clearly understood. The purpose of this paper is to explain why there may be too much R&D. According to conventional wisdom, over-investment is caused by a business-stealing effect that arises as the latest innovator destroys and/or appropriates previous incumbent's rents. We argue that in standard models, business stealing by itself cannot result in excessive R&D. The conventional wisdom is therefore misleading, over-estimating the possibility that the market may be biased towards excessive R&D. We explain the other effects that must be at work for over-investment to be possible, correcting several conjectures made in the literature.

While the term "business stealing" generally refers to the effect that entry by the latest innovator has on the profits of the previous incumbent, its exact meaning is somewhat ambiguous. Some authors view business stealing simply as the loss to the previous monopolist resulting from the latest innovation. The claim is that over-investment is due to this pecuniary externality not being internalised by the latest innovator. Other authors develop this idea by viewing business stealing as a redistribution of rents from past innovators to the latest. By appropriating the previous incumbent's rents – so the argument goes – the latest innovator obtains more than the social value of his innovation, and therefore has an excessive incentive to invest in R&D.

We start by explaining graphically why, in standard quality-ladder models, the cause of over-investment in R&D cannot be business stealing, whatever its exact interpretation. Consider the market structure common to those models, characterised by perpetual leapfrogging and price competition among successive innovators. Figure 1 shows the product market equilibrium in the case of cost-reducing innovations, which is equivalent to the case of quality-improving innovations if goods are measured in efficiency units.

The figure shows that when a new innovation arrives, the previous incumbent's profits do not disappear but are turned into consumer surplus. This means that in the social welfare calculation, a positive externality offsets a negative one. In fact, the increase in consumer surplus caused by the latest innovation is always larger than the past incumbent's profits. In other words, there is a consumer surplus effect that always prevails over the business-stealing effect. The figure also shows that not a penny of the past incumbent's profits is gathered by the latest innovator. The latest innovator extracts his profit exclusively from the new value he creates for society. Thus, there is no redistribution of rents from past innovators to current ones.[1,2] This

[1] For a similar graphical analysis see Stoneman (2005), who also concludes that the private gain from an innovation cannot exceed the social gain.

[2] In the remainder of the paper, the term "business stealing" shall therefore be used to refer to the destruction of the past incumbent's rents.

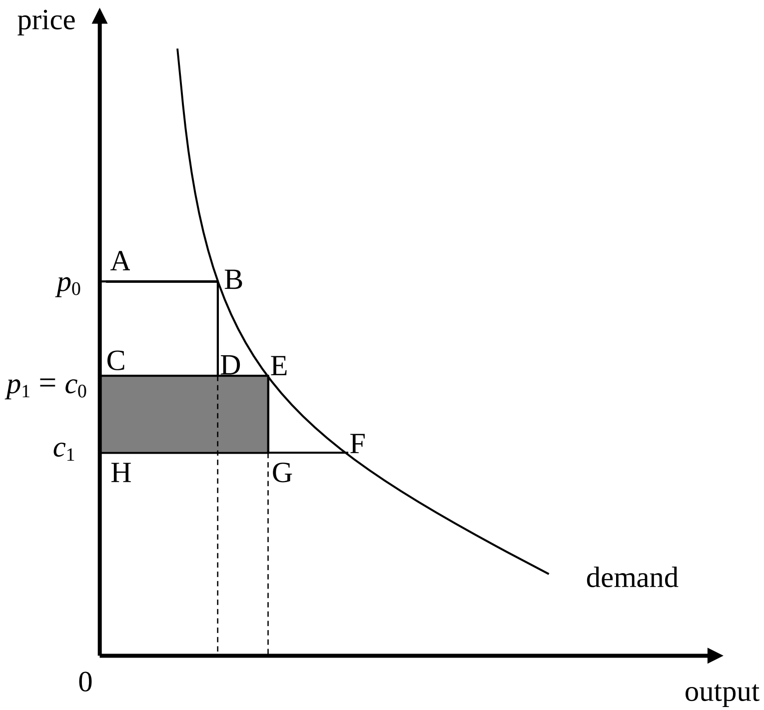

Fig. 1(a). Non-drastic innovations

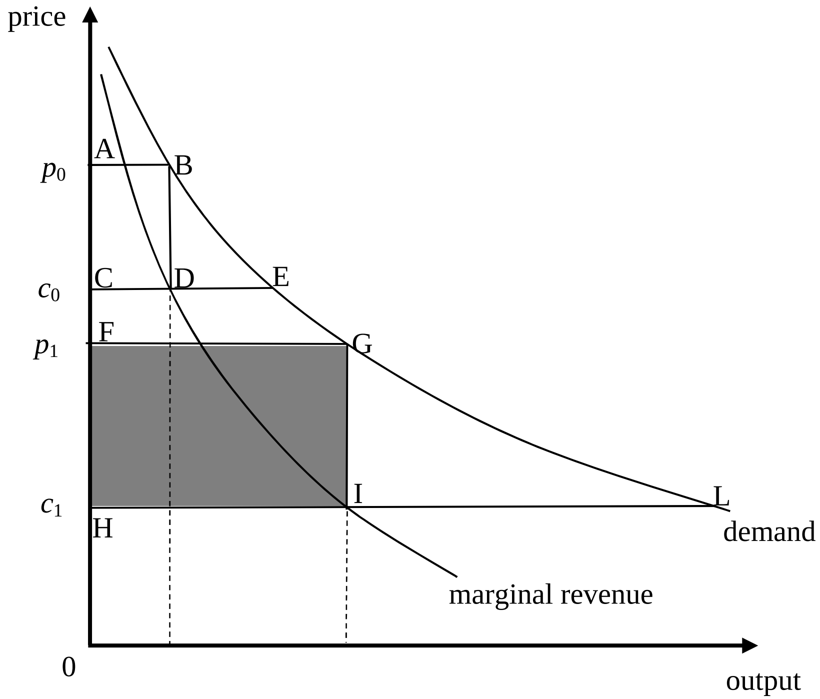

Fig. 1(b). Drastic innovations

The innovation reduces the unit production cost from c_0 to c_1. If the innovation is non-drastic, as in this panel, the latest innovator sets a limit price $p_1 = c_0$. The area ABCD represents the past incumbent's profits; the grey area CEGH, the latest innovator's ones. The increase in consumer surplus, which is the area ABEC, exceeds the profit loss to the previous incumbent by the area BDE. The latest innovator does not obtain any of the profits of the previous incumbent; he obtains a share of the new value he has created (the area CEFH).

This panel shows the case of drastic innovations. In this case, the latest innovator sets the monopoly price p_1. Again, the area ABCD represents the past incumbent's profits; the grey area FGHI, the latest innovator's ones. The increase in consumer surplus now exceeds the profit loss to the previous incumbent by the sum of the areas BDE and CEFG. As in the case of non-drastic innovations, the latest innovator does not obtain any of the profits of the past incumbent; he obtains only a share of the new value he has created (which is now the area CELH).

implies that the social value of an innovation is always greater than its private value.[3]

We conclude that there is no reason for over-investment in R&D here.[4] However, a number of Schumpeterian models, starting from the seminal contributions of Aghion and Howitt (1992) and Grossman and Helpman (1991a), have found that either under- or over-investment in R&D may occur in the market equilibrium. If, as we contend, business stealing cannot be held responsible for this result, then other effects must.

One obvious candidate is the R&D congestion effect (also known in the literature as the crowding effect, the stepping-on-toes effect, or the winner-takes-all effect). This is represented by the decrease in the probability of a firm's competitors' success resulting from an increase in that firm's R&D investment, or the corresponding increase in their cost needed to achieve the innovation with a given probability. That this negative externality can generate excessive R&D is clearly understood. However, the R&D congestion effect vanishes when the returns to R&D are constant, as in many quality-ladder models, including Aghion and Howitt (1992) and Grossman and Helpman (1991a).

With constant returns to R&D, the only cause of over-investment is a monopoly distortion effect noted by Aghion and Howitt (1992), but hardly mentioned in subsequent studies. This is an equilibrium property of so-called "scarce-factors" models, where there is a factor of production (labour) that is in fixed supply and can be used exclusively for the production of innovative goods or for the purposes of research.

The monopoly distortion effect can be described as follows. Since innovators have market power, the markets for innovative goods are imperfectly competitive. This means that the wage rate is lower than the marginal productivity of labour. As a result, firms that hire labour to conduct their research are faced with a price (the wage rate) that is lower than the true social cost of labour (its marginal productivity), and so have an excessive incentive to invest in R&D. Mechanically speaking, the monopoly distortion means that the production of innovative goods is inefficiently low and too little labour is employed in the innovative goods sector. Since labour supply is fixed, and the only alternative use of labour (in the model) is to conduct R&D, then excessive labour must be employed on R&D.

The monopoly distortion effect disappears in "lab-equipment" models, where the R&D input is the final good rather than labour. We therefore submit that these models cannot exhibit over-investment if there are constant returns to R&D. The

[3] To be precise, the argument shows that the *flow* social value of an innovation is greater than the *flow* private value. However, while social value is a permanent addition to social welfare, the innovator's profits are terminated by the arrival of the next innovation. Thus, the gap between total *discounted* social and private gains is even larger.

[4] The figure shows the case of constant marginal costs, i.e. the assumption commonly made in the endogenous growth literature. However, the same conclusion would hold with increasing marginal costs, provided that the size of the innovation is sufficiently large. For smaller innovations, or with decreasing marginal costs, the equilibrium may involve mixed strategies, which would complicate the analysis.

first section endeavours to prove these claims for quality ladder models, whereas section 2 examines models with expanding product variety.

1. Quality-ladder Models

In this section, we argue that if one eliminates the monopoly distortion effect from quality-ladder models of endogenous growth, and there are no R&D congestion effects, then there cannot be any over-investment in R&D. We focus on the seminal contributions by Aghion and Howitt (1992) and Grossman and Helpman (1991a). However, the same comments would apply to the many models that have extended earlier theories in various ways, most notably those models with no scale effects (e.g. Howitt, 1999), and those in which the technological leader may innovate repeatedly (e.g. Grossman and Helpman, 1991b). In the main text, we shall use the Aghion and Howitt (1992) model for the purposes of our demonstration. Appendix A shows that the same conclusions hold for the model proposed by Grossman and Helpman (1991a).

We consider two ways of removing the monopoly distortion effect. First, we retain the "scarce-factors" framework but assume that labour used in research is taxed at a rate equal to the mark-up charged by innovators in the innovative goods sector. With such a tax in place, the labour cost faced by research firms will equal the true social cost. Secondly, we reformulate the model in a "lab-equipment" framework, where the R&D input is the final good rather than labour. Within this framework, monopoly distortions no longer affect the cost of conducting the research, and so they cannot generate excessive R&D.

1.1. *Model Assumptions*

In order to make the paper self-contained, we first provide a brief account of the assumptions of Aghion and Howitt (1992) and the derivation of the equilibrium.

There are three goods in the economy: labour, an intermediate good, and a final good. Labour can be employed in the production of the intermediate good or in research, the intermediate good is used to produce the final good, and the final good is consumed. The economy is populated by L identical, infinitely-lived households. Each household inelastically supplies one unit of labour and maximises the discounted utility $u = \int_0^\infty c(\tau)e^{-r\tau}d\tau$, where r is the rate of time preference and $c(\tau)$ the per capita consumption of the final good. With a linear instantaneous utility, the interest rate is directly given by the rate of time preference r.

The quality of the intermediate good increases over time due to technical progress. Each innovation improves said quality by a constant factor. Innovations occur at random time intervals according to a Poisson process with a hazard rate that depends on the rate of R&D investment. We refer to period t as the time interval between innovation t and innovation $t+1$. The economy is stationary within each period, but it jumps up by a constant factor from one period to the next.

One unit of labour is required to produce one unit of the intermediate good, regardless of the latter's vintage. Normalising to one the quality of the intermediate good at time zero, and assuming that only the latest vintage is produced, the production function of the final good in period t is:

$$y_t = \gamma^t x_t^\alpha \qquad \text{with } 0 < \alpha < 1, \tag{1}$$

where y_t is the output of the final good, x_t is the input of the intermediate good, and $\gamma^{\frac{1}{\alpha}}$ is the size of quality improvements.

To discover higher quality products, firms engage in R&D races. There is free entry into each race, and all firms have the same R&D technology. If n_t units of labour are used in research in period t, the new highest quality product $t+1$ is discovered with an instantaneous probability λn_t. The parameter λ is the productivity of labour in research. The winner of a R&D race becomes the sole producer of the highest quality product.

1.2. *Equilibrium*

Standard arguments imply that in equilibrium incumbents do not participate in the race for the subsequent innovation, and so there is systematic leapfrogging. Profit maximisation by the final good sector leads to the following inverse demand for the intermediate good:

$$p_t = \gamma^t \alpha x_t^{\alpha-1} \tag{2}$$

where p_t is the price. The elasticity of demand is $\frac{1}{1-\alpha}$. Each vintage of the intermediate good has a constant marginal cost equal to the wage rate w_t. This implies that the monopoly price is $p_t = \dfrac{w_t}{\alpha}$, and that innovations are drastic if $\gamma^{\frac{1}{\alpha}} \geq \dfrac{1}{\alpha}$. In the main text we focus on the case of drastic innovations; the case of non-drastic innovations leads to the same results and is dealt with in footnotes.[5]

The equilibrium wage rate is

$$w_t = \gamma^t \alpha^2 x_t^{\alpha-1}, \tag{3}$$

and innovator t earns a flow of profit equal to

$$\pi_t = \gamma^t \alpha (1-\alpha) x_t^\alpha.\text{[6]} \tag{4}$$

[5] When innovations are non-drastic, i.e. when $\gamma^{\frac{1}{\alpha}} < \dfrac{1}{\alpha}$, the latest innovator engages in limit pricing. Measuring the intermediate good in efficiency units relative to the last vintage, the unit cost of vintage $t-1$ (the latest innovator's most efficient competitor) is $\gamma^{\frac{1}{\alpha}} w_t$. The limit price is then $p_t = \gamma^{\frac{1}{\alpha}} w_t$.

[6] With non-drastic innovations, the equilibrium wage rate is

$$w_t = \gamma^{t-\frac{1}{\alpha}} \alpha x_t^{\alpha-1},$$

and the innovator's flow profit is

$$\pi_t = \gamma^t \left(1 - \gamma^{-\frac{1}{\alpha}}\right) \alpha x_t^\alpha.$$

To obtain the value of innovation t, this profit flow must be discounted by the interest rate r augmented by the probability λn_t that the next innovation arrives. Therefore, the value of innovation t is

$$v_t = \frac{\pi_t}{r + \lambda n_t}. \tag{5}$$

In period t firms race to discover innovation $t+1$. The free-entry condition in patent races requires that v_{t+1} be equal to the unit cost of R&D, i.e.,

$$v_{t+1} = \frac{w_t}{\lambda}. \tag{6}$$

Although the model admits also non-stationary equilibria, we focus on steady states where $n_{t+1} = n_t$. Combining (3)-(6) and using the labour market clearing condition $x_t = L - n_t$, one finally obtains the following market equilibrium condition:

$$\frac{\gamma^{t+1}\alpha(1-\alpha)(L-n^*)^\alpha}{r+\lambda n^*} = \frac{\gamma^t \alpha^2 (L-n^*)^{\alpha-1}}{\lambda}. \tag{7}$$ [7]

The left-hand side of (7) is the marginal private benefit of R&D, and the right-hand side is the marginal private cost, i.e. the wage rate divided by the productivity of labour employed in R&D.

To ascertain whether there is over- or under-investment in R&D in the market equilibrium, we evaluate the sign of the change in social welfare u associated with a small permanent increase in the rate of innovative activity n. In a steady state, the expected discounted utility is

$$u = \frac{(L-n)^\alpha}{r - \lambda n(\gamma - 1)}. \tag{8}$$

Since the model does not possess any transitional dynamics, the change in social welfare associated with a small permanent increase in n is simply $\frac{du}{dn}$. We have:

$$\frac{du}{dn} \propto \frac{(\gamma-1)\gamma^t (L-n)^\alpha}{r - \lambda n(\gamma-1)} - \frac{\gamma^t \alpha (L-n)^{\alpha-1}}{\lambda}, \tag{9}$$

where the symbol $\propto$ means "has the same sign as." The first term on the right-hand side is the discounted value of the increase in output resulting from innovation $t+1$, $(\gamma^{t+1} - \gamma^t)(L-n)^\alpha$, which is the marginal social gain from said innovation.[8] The second term is the foregone current output when one unit of labour is used in research, i.e. the marginal social cost of R&D.

At the market equilibrium, (9) becomes

$$\frac{du}{dn}\Big|_{n=n^*} \propto \frac{(\gamma-1)}{r - \lambda n^*(\gamma-1)} - \frac{\gamma(1-\alpha)}{r+\lambda n^*}. \tag{10}$$

[7] With non-drastic innovations, the corresponding formula is

$$\frac{\gamma^{t+1}\left(1-\gamma^{-\frac{1}{\alpha}}\right)\alpha(L-n^*)^\alpha}{r+\lambda n^*} = \frac{\gamma^{t-\frac{1}{\alpha}}\alpha(L-n^*)^{\alpha-1}}{\lambda}.$$

[8] This is always positive since $r > \lambda n(\gamma - 1)$ by the transversality condition.

1.3. *Removing the Monopoly Distortion Effect*

As Aghion and Howitt (1992) pointed out, the derivative (10) can be either positive or negative, meaning that in the market equilibrium there may be too little, or too much, R&D. The reason for this is that there are various differences between the marginal private and social costs and benefits of conducting the research. One is that the private cost of conducting the research (i.e., the right-hand side of (7)) is lower than the social cost (i.e., the second term of (9)) by a factor of α. This difference is due to the monopoly distortion effect: because of monopoly pricing, the marginal productivity of labour used in the production of the intermediate good is greater than the wage rate by a factor of α. As a result, firms that hire labour to conduct R&D are faced with a labour cost that is lower than the true social cost.

We submit that this monopoly distortion effect is the only possible reason for over-investment in R&D in this model. To prove this, we remove the effect and show that, as a result, the derivative (10) is always positive.

1.3.1. *A corrective tax on labour employed in R&D*

One way to remove the monopoly distortion effect is to impose a corrective tax. To be precise, labour employed in R&D must be taxed at rate θ equal to the mark-up charged by innovators in the innovative goods sector, i.e. $\theta = \frac{1}{\alpha} - 1$ (fiscal revenue is then returned to consumers as a lump-sum subsidy.) This guarantees that the labour cost perceived by R&D firms is equal to the true social cost of labour.

With this tax in place, the free entry condition in patent races becomes $v_{t+1} = \frac{w_t}{\alpha\lambda}$, and the equilibrium rate of innovation is given by:

$$\frac{\gamma^{t+1}\alpha(1-\alpha)(L-n^*)^{\alpha}}{r+\lambda n^*} = \frac{\gamma^t \alpha (L-n^*)^{\alpha-1}}{\lambda}.\text{[9]} \tag{11}$$

The only difference with (7), i.e. the equilibrium condition of the original model, is that the right-hand side is now divided by a factor of α. This is how Aghion and Howitt themselves identify the monopoly distortion effect (see Aghion and Howitt, 1992, p. 338).

Although the private cost of conducting the research now coincides with the social cost, there are still differences between its private and social benefits. First, the private benefit flow from the innovation (i.e., the numerator of the left-hand side of (11)) is different from the social benefit flow (i.e., the numerator of the first term of

[9] With non-drastic innovations, the corrective tax is $\theta = \gamma^{\frac{1}{\alpha}} - 1$, and the modified equilibrium condition is

$$\frac{\gamma^{t+1}\left(1-\gamma^{-\frac{1}{\alpha}}\right)\alpha(L-n^*)^{\alpha}}{r+\lambda n^*} = \frac{\gamma^t \alpha (L-n^*)^{\alpha-1}}{\lambda}.$$

(9)). Aghion and Howitt (1992) break this difference down into two separate effects, the "consumer surplus" and the "business-stealing" effects, which are of opposite sign. However, we show below that the total effect can be signed unambiguously. The social benefit flow from the innovation, that is to say, is always greater than the private benefit flow, which is in keeping with our graphical analysis set out in the introduction.

Second, the private benefit flow from the innovation is discounted more heavily than the corresponding social benefit flow. This reflects the fact that the social benefit is permanent, whereas the innovator's profits end when the next innovation arises, and that in the market equilibrium the innovator is not rewarded for opening the way to subsequent improvements.

These observations lead us to formulate the following:

Proposition 1 *In the modified Aghion and Howitt model, where the monopoly distortion effect is removed by a corrective tax on labour used in research, there is always under-investment in R&D.*

Proof. Plugging (11) into (9) we obtain

$$\frac{du}{dn}\mid_{n=n^*} \propto \frac{(\gamma-1)}{r-\lambda n^*(\gamma-1)} - \frac{\gamma\alpha(1-\alpha)}{r+\lambda n^*},$$

Since the denominator of the first term is always smaller than that of the second term, a sufficient condition for $\frac{du}{dn}\mid_{n=n^*}$ to be positive is that:

$$1-\frac{1}{\gamma} \geq \alpha(1-\alpha).$$

Notice that the left-hand side of this inequality is increasing in γ, and the condition for innovations to be drastic is $\gamma^{\frac{1}{\alpha}} \geq \frac{1}{\alpha}$. It follows that a sufficient condition for $\frac{du}{dn}\mid_{n=n^*}$ to be positive is that the above inequality holds at $\gamma^{\frac{1}{\alpha}} = \frac{1}{\alpha}$, that is

$$1-\alpha^{\alpha}-\alpha(1-\alpha) \geq 0.$$[10]

Simple algebra shows that this condition indeed holds as an equality for $\alpha = 0$ and $\alpha = 1$, and as a strict inequality for $0 < \alpha < 1$. ■

[10] With non-drastic innovations, the relevant condition is $(\gamma-1)\frac{1}{\alpha} \geq \gamma\left(1-\gamma^{-\frac{1}{\alpha}}\right)$, or

$$\Gamma(\gamma,\alpha) \equiv \frac{\gamma-1}{\gamma-\gamma^{1-\frac{1}{\alpha}}} \geq \alpha.$$

Since $\lim_{\gamma\to 1}\Gamma(\gamma,\alpha) = \alpha$, we just need to show that $\Gamma(\gamma,\alpha)$ is non-decreasing in γ on $[1,\alpha^{-\alpha}]$. We have:

$$\frac{\partial\Gamma}{\partial\gamma}(\gamma,\alpha) = \frac{1}{\alpha\gamma^{\frac{1}{\alpha}}\left(\gamma-\gamma^{1-\frac{1}{\alpha}}\right)^2}S(\gamma,\alpha),$$

where $S(\gamma,\alpha) \equiv \left(1+\alpha\gamma^{\frac{1}{\alpha}}-\gamma-\alpha\right)$. Clearly, $S(1,\alpha) = 0$ and $\frac{\partial S}{\partial\gamma}(\gamma,\alpha) = \gamma^{\frac{1}{\alpha}-1}-1 \geq 0$, which proves the result.

The same result is obtained if instead of correcting the monopoly distortion by imposing a tax on labour employed in R&D, one subsidises the production of the intermediate good. The appropriate subsidy rate is $s = \frac{1}{\alpha} - 1$, and the "income effects" of the subsidy would be sterilised by a tax on the innovator's profits at rate $1 - \alpha$. One immediately sees that this would lead to the same equilibrium as that of the corrective tax on labour.

1.3.2. *A lab-equipment model*

Another way of removing the monopoly distortion effect is to use a lab-equipment reformulation of the model. Originally proposed by Rivera-Batiz and Romer (1991) in a model of expanding product variety, the lab-equipment formulation assumes that the R&D input is the final good rather than labour. Thus, monopoly pricing distortions no longer affect the cost of conducting the research.

The lab-equipment formulation has since been adopted in many quality-ladder models as well. It is our contention that these lab-equipments models must generate too little R&D when there are constant returns to R&D. Appendix B verifies this claim for two of the most popular lab-equipment models proposed in the quality-ladder literature.[11] However, both models modify various other assumptions made by Aghion and Howitt (1992), allowing for a strictly concave instantaneous utility function and many intermediate goods. In order to show that these other changes are not responsible for the result, we develop a simple lab-equipment model that departs from the original Aghion and Howitt model only insofar as it assumes that the final good, rather than labour, is the input used to carry out R&D.

The variable n_t now denotes the amount of the final good used for the purposes of research at time t, and the labour market clearing condition is simply $x_t = L$. The instantaneous probability of discovery is $\lambda_t n_t$. To guarantee the existence of a steady state, we must now assume that the productivity of the R&D input decreases over time. The reason for this is that in a steady state the R&D investment n_t must grow by a factor of γ from one period to the next. In order for the hazard rate $\lambda_t n_t$ to be constant, the R&D productivity λ_t must then fall at a rate of γ. This requires the assumption $\lambda_t = \lambda \gamma^{-t}$.[12]

The free-entry condition now becomes $v_{t+1} = \frac{\gamma^t}{\lambda}$, and the market equilibrium is given by

$$\frac{\gamma^{t+1}\alpha(1-\alpha)L^{\alpha}}{r + \lambda_{t+1}n_{t+1}} = \frac{\gamma^t}{\lambda}. \tag{12}$$

[11] These are the textbook models of Barro and Sala-i-Martin (2004) and Acemoglu (2009). Sharing the widely-held, albeit mistaken, belief that over-investment is due to the business stealing effect, both textbooks claim that over-investment is possible. However, Appendix B proves that this is not the case.

[12] For similar knife-edge assumptions in lab-equipment models see, for example, Barro and Sala-i-Martin (2004) and Acemoglu (2009). When labour is the R&D input, the cost of R&D increases automatically at the appropriate rate, since the wage rate increases as the economy grows.

Focusing on steady states where $n_t = n\gamma^t$, discounted social welfare is

$$u = \frac{L^\alpha - n}{r - \lambda n(\gamma - 1)}. \quad (13)$$

Proceeding as before, we get:

Proposition 2 *In the lab-equipment version of the Aghion and Howitt model, there is always under-investment in R&D.*

Proof. The market equilibrium rate of innovation is positive if, and only if,

$$\gamma\alpha(1-\alpha) > \frac{r}{\lambda}L^{-\alpha}.$$

The marginal effect on social welfare of an increase in the rate of innovation is now

$$\frac{du}{dn}\;|_{n=n^*} \propto -\frac{r}{\lambda}L^{-\alpha} + (\gamma - 1).$$

A sufficient condition for $\frac{du}{dn}\;|_{n=n^*}$ to be positive when $n^* > 0$ is that:

$$1 - \frac{1}{\gamma} \geq \alpha(1-\alpha).$$

This is exactly the same condition found in the proof of Proposition 1, and we already know that it is always satisfied. ■

As we have argued in the introduction, the result follows from the fact that with price competition and vertically-differentiated products, the social value of an innovation is always greater than its private value. When a new innovation emerges, the increase in consumer surplus always exceeds the loss to the previous incumbent, and there is no redistribution of rents from past to current innovators.

2. Models with Expanding Product Variety

We now turn to models of endogenous growth with expanding product variety. We argue that the business stealing effect alone cannot result in over-investment in these models either. Unlike Schumpeterian models, however, not even the monopoly distortion effect suffices to generate excessive R&D: over-investment is only possible in the presence of R&D congestion effects.

These claims may sound surprising at first. The industrial organisation literature shows that a rent-shifting effect can indeed occur when goods are horizontally differentiated, and that, as a result, the private value of innovations may be grater than their social value. This suggests that models with expanding product variety are more likely to generate over-investment in R&D than quality-ladder models.

However, the standard specification of technology (or preferences) in expanding-variety models precludes the possibility of over-investment in the absence of congestion effects. The reason for this is simple. Consider the standard Dixit and Stiglitz

(1977) production function

$$y = \left(\sum_{i=1}^{n} d_i^{\alpha} \right)^{\frac{1}{\alpha}} \tag{14}$$

where α is a parameter lower than one, n now denotes the number of varieties already invented and d_i the quantity of variety i. With this technology, the equilibrium price for each variety is $p = w/\alpha$, with a constant mark-up of $\frac{1-\alpha}{\alpha}$. This implies that the innovator's profit is[13]

$$\pi = (1-\alpha)\,\frac{y}{n}. \tag{15}$$

By increasing the number of existing varieties, innovation decreases the profits of the past incumbents.[14] However, the productivity gain created by the invention of a new variety is:

$$\frac{dy}{dn} = \frac{1-\alpha}{\alpha}\frac{y}{n}. \tag{16}$$

Clearly, with a Dixit-Stiglitz production function, the innovator's profit (i.e. the private value of an innovation) is just a share $\alpha < 1$ of the productivity gain created by the innovation (its social value).[15]

The implications of this depend on what other assumptions are made. In scarce-factors models, where the monopoly distortion effect is at work, the wage rate is a fraction α of the marginal productivity of labour. Therefore, research firms perceive a labour cost that is a fraction α of the true social cost of labour. On the other hand however, as we have just seen, the private value of an innovation is a share α of its social value. This means that the monopoly distortion effect exactly compensates the fact that innovators do not capture the social value of innovations fully. In the absence of other effects, the rate of innovation in the market equilibrium would be just optimal. However, in expanding variety models, sustained growth can only be guaranteed by assuming that the invention of a new variety facilitates future

[13] The demand function for variety i is

$$p_i = \left(\frac{y}{d_i}\right)^{1-\alpha}.$$

At the equilibrium price $p = w/\alpha$, the innovator's profit is

$$\pi = (1-\alpha)y\left(\frac{w}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}.$$

In a steady state in which the share of labour used in the production of intermediate good is constant, we have

$$\frac{w}{\alpha} = n^{\frac{1-\alpha}{\alpha}}$$

from which the expression used in the text follows.

[14] This is not generally true in models with expanding product variety. The specification adopted by Romer (1990) and Aghion and Howitt (2009), for instance, implies that the introduction of new varieties does not affect the profits of past incumbents. In these models, π is independent of n.

[15] Equations (15) and (16) give the flow value of the innovator's profits and the productivity gain, respectively. In expanding variety models, both flows are permanent, and so the ratio between the flows equals the ratio of the respective discounted values.

innovation. This positive externality is another source of under-investment in R&D, which means that the equilibrium rate of innovation must be inefficiently low.

In lab-equipment models, things are simpler. The fact that the monopoly distortion effect vanishes immediately implies that over-investment in R&D is not possible in the absence of R&D congestion effects. Therefore, both in lab-equipment and in scarce-factors models of expanding product variety with a Dixit-Stiglitz specification of the technology, there can be no over-investment.

A few authors have tried to overcome this conclusion. Benassy (1998) replaces the Dixit-Stiglitz production function (14) with

$$y = n^{\eta} \left(\sum_{\omega=1}^{n} d_{\omega}^{\alpha} \right)^{\frac{1}{\alpha}} . \tag{17}$$

Using this alternative formulation, he shows that over-investment in R&D is possible when $\eta < 0$. However, when the parameter η is negative the invention of a new variety shifts a part of the economy's production possibility frontier down.[16] This means that the invention of a new good entails a kind of technological *regression*. If innovation can only move the economy's production possibility frontier up, the additional parameter η cannot be negative. This reinstates the under-investment outcome.

Jones and Williams (2000) assume that the invention of new goods occurs in clusters, which include both new varieties and alternative versions of existing varieties (which they refer to as "upgrades"). The "upgrades" are not really any more productive than existing varieties; rather, they are perfect substitutes. However, innovators can bundle upgrades and new varieties, which allows them to extract a greater share of the social value of the innovation at the expense of previous incumbents. That is, bundling magnifies the business stealing effect. Nevertheless, Appendix C shows that in the Jones and Williams model the possibility of over-investment in R&D is entirely due to the R&D congestion effect. In other words, with constant returns to R&D there will always be under-investment in research.[17] Appendix C also provides an intuitive explanation for this result, by showing that the business stealing effect is again always dominated by the consumer surplus effect.

3. Conclusion

[16] This can be easily seen by assuming that the input of the newly invented variety is nil, while those of old varieties are the same as before the innovation: with the production function (17), output would then be lower after the innovation if $\eta < 0$.

[17] In their calibration of the model, Jones and Williams (2000) do find that the consumer surplus effect is several times larger than the business stealing effect. However, they claim that in theory the latter effect could be stronger than the former.

This paper aims to dispel the widespread belief that business stealing by itself may cause over-investment in R&D in models of endogenous growth. We have shown that in standard quality-ladder models with constant returns to R&D, the source of over-investment is a general equilibrium, monopoly distortion effect that arises only in scarce-factor models and is, arguably, an artefact of special modelling assumptions. In models with expanding variety, not even this monopoly distortion effect suffices to generate over-investment.

Our analysis therefore implies that the only robust cause of over-investment in R&D is the congestion effect that arises when the returns to R&D are decreasing. In this case, if each innovator is small in relation the aggregate, then it will fail to internalise the negative externality that its R&D investment imposes on others, leading to excessive investment in R&D (Stokey, 1995).

There are several reasons why a clear understanding of the possible sources of over-investment is important. Firstly, it provides scholars with simple guidelines as to whether any particular model may exhibit excessive R&D or not. Secondly, it may help in empirically evaluating whether there is too much or too little R&D investment in real life. The key parameter appears to be the degree of decreasing returns to R&D, which determines the magnitude of the R&D congestion effects. Empirical estimates suggests that the elasticity of the "innovation production function" is around 0.5, thus indicating that returns to scale may be significantly declining.[18] Finally, our analysis may clarify the policy implications of the theory. Consider, for instance, the debate about the subsidisation of large or small innovations (Segerstrom, 1998; Li, 2003). The models used in this debate are a development of Grossman and Helpman (1991a). Thus, they exhibit a monopoly distortion effect which increases as the size of the innovation increases. To the extent that the monopoly distortion effect is a model artefact, the said models may underestimate the need to subsidise large innovations rather than smaller ones.

[18] For a survey of such empirical studies see, for example, Scotchmer (2004).

References

[1] Acemoglu, D. (2009). *Introduction to modern economic growth*, Princeton: Princeton University Press.

[2] Aghion, P. and Howitt, P. (1992). 'A model of growth through creative destruction', *Econometrica*, vol. 60(2), pp. 323-351.

[3] Aghion, P. and Howitt, P.(2009). *The economics of growth*, Cambridge: MIT Press.

[4] Barro, R. J. and Sala-i-Martin, X. (2004). *Economic growth*, Cambridge: MIT Press.

[5] Benassy, J. P. (1998). 'Is there always too little research in endogenous growth with expanding product variety?', *European Economic Review*, vol. 42(1), pp. 61-69.

[6] Burstein, M.L. (1960). 'The economics of tie-in sales', *Review of Economics and Statistics*, vol. 42(1), pp. 68-73

[7] Dixit, A. and Stiglitz, J. (1977). 'Monopolistic competition and optimum product diversity', *American Economic Review*, vol. 67(3), pp. 297-308.

[8] Grossman, G. M. and Helpman, E. (1991a). 'Quality ladders in the theory of growth', *Review of Economic Studies*, vol. 58(1), pp. 43–61.

[9] Grossman, G. M. and Helpman, E. (1991b). 'Quality ladders and product cycles', *Quarterly Journal of Economics*, vol. 106(2), pp. 557-586.

[10] Jones C. and Williams, J. (2000). 'Too much of a good thing? The economics of investment in R&D', *Journal of Economic Growth*, vol. 5(1), pp. 65-85.

[11] Li, C. W. (2003). 'Endogenous growth without scale effects: a comment', *American Economic Review*, vol. 93(3), pp. 1009-1017.

[12] Segerstrom, P. S. (1998). 'Endogenous growth without scale effects', *American Economic Review*, vol. 88(5), pp. 1290-1310.

[13] Stoneman, P. (2005). 'Creative destruction and unverinvestment in R&D: did Arrow point the wrong way?', paper presented at the 2005 John Barber Conference, Manchester.

[14] Rivera-Batiz, L. and Romer, P. M. (1991). 'Economic integration and endogenous growth', *Quarterly Journal of Economics*, vol. 106(2), pp. 531-555.

[15] Romer, P. (1990). 'Endogenous technological change', *Journal of Political Economy*, vol. 98(1), pp. 71-102.

[16] Scotchmer, S. (2004). *Innovation and incentives*, Cambridge: MIT Press.

[17] Stokey, N. (1995). 'R&D and economic growth', *Review of Economic Studies*, vol. 62(3), pp. 469-489.

Appendix A

This Appendix shows that the monopoly distortion effect is the only cause of over-investment in R&D in the model of Grossman and Helpman (1991a), too. We again eliminate the monopoly distortion effect either by assuming that labour employed on R&D is taxed at a rate equal to the mark-up charged by innovators in the innovative goods sector, or by reformulating the model as a lab-equipment model.

We start from the lab-equipment version of Grossman and Helpman (1991a). We depart from the original model only insofar as the final good, rather than labour, is the R&D input.[19] To facilitate the comparison, we use the same notation as Grossman and Helpman (1991a).

There are three types goods: a final good, labour, and a continuum of intermediate goods indexed by $\omega \in [0, 1]$. The quality of intermediate goods increases over time due to technical progress. Normalise the quality of each good at time zero to 1, and denote by $j(\omega, t)$ the number of innovations achieved in sector ω by time t. Thus, the highest quality of good ω which can be produced at time t is $\lambda^{j(\omega,t)}$, where $\lambda > 1$ denotes the size of each innovation.

Regardless of quality j and variety ω, one unit of labour is required to produce one unit of intermediate good. Consequently, each firm has a constant marginal cost of production equal to the wage rate w. In each industry, successive innovators compete in prices. As a result, in equilibrium only the state-of-the-art version of the good is produced in each industry.

The final good can be consumed or used as an input for R&D. This good is taken as the numeraire. It is produced in a perfectly competitive market using intermediate goods only. The production function is:

$$\log y(t) = \int_0^1 \log \left[\lambda^{j(\omega,t)} d(j, \omega, t) \right] d\omega,$$

where $y(t)$ is the output of the final good and $d(j, \omega, t)$ denotes the input of intermediate good ω of vintage j at time t. Profit maximisation by the final good sector implies a unit-elastic demand for the intermediate good:

$$d(j, \omega, t) = \frac{y(t)}{p(j, \omega, t)}$$

where $p(j, \omega, t)$ is the price.

The economy is populated by L identical, infinitely-lived households. Each household inelastically supplies one unit of labour and maximises the discounted utility:

$$u = \int_0^\infty \log c(t) e^{-\rho t} dt,$$

[19] For expositional reasons, we follow the "intermediate goods" interpretation of the model and choose a different numeraire from Grossman and Helpman, but these changes do not affect the substantive conclusions.

where ρ is the rate of time preference and $c(t)$ is the per capita consumption of the final good.

To discover higher quality products, firms in each industry engage in R&D races. There is free entry into each R&D race. All firms have the same R&D technology.

Here we depart from Grossman and Helpman (1991a) by assuming that the final good, rather than labour, is the only input used to do R&D. Any R&D firm i that uses $n_i(j, \omega, t)$ units of the final good at time t may discover the next higher quality product $j+1$ in industry ω with instantaneous probability $\frac{n_i(j,\omega,t)}{a(t)}$. The variable $a(t) > 0$ is an index of R&D cost. The returns to engaging in R&D races are independently distributed across firms, across industries, and over time. Thus, the industrywide instantaneous probability of success at time t is simply $\iota(j, \omega, t) = \frac{n(j,\omega,t)}{a(t)}$, where $n = \sum_i n_{i,}$ is the industrywide investment in R&D. Given the symmetric structure of the model, we focus on equilibrium behaviour where the R&D intensity $n(j, \omega, t)$ is the same in all industries ω at time t, $n(t)$.

As in Grossman and Helpman's original model, the winner of a R&D race becomes the sole producer of the highest quality product. Standard arguments show that in equilibrium incumbents do not participate in the race for the subsequent innovation, so there is systematic leapfrogging.

We are interested in a steady state where output, consumption and R&D investment grow at a common and constant rate of g. Standard calculations show that the growth rate is $g = \iota \log \lambda$. To guarantee the existence of such a steady state, we must ensure that the rate of innovation can be constant. Since in a steady state R&D investment $n(t)$ grows at rate of g then in order for the hazard rate $\iota = \frac{n(t)}{a(t)}$ to be constant the unit cost of R&D, $a(t)$, must grow at rate of g. This requires the assumption $a(t) = ae^{gt}$.

Price competition between the latest innovator and the previous incumbent leads to a limit pricing equilibrium where the quality leader sets the price $p = \lambda w$, earning the profit flow

$$\pi(t) = \frac{\lambda - 1}{\lambda} y(t).$$

The intertemporal maximisation problem of the representative household yields the well-known Euler equation

$$\frac{\dot{c}(t)}{c(t)} = r(t) - \rho.$$

A standard no-arbitrage condition requires that

$$rv(t) = \pi(t) + \dot{v}(t) - \iota v(t),$$

where $v(t)$ denotes the value of the leading firms. Finally, the free-entry condition requires that the value of conducting R&D is equal to the unit R&D cost, i.e.,

$$v(t) = a(t).$$

Combining all these equilibrium conditions we get

$$a(t) = \frac{\frac{\lambda-1}{\lambda} y(t)}{\rho + \iota}.$$

Finally, using the fact that in a steady state $y(t) = Le^{gt}$, we can solve the above equation to obtain the steady state rate of innovation

$$\iota^* = \frac{L}{a}\frac{\lambda - 1}{\lambda} - \rho.$$

This formula is identical to the one derived by Grossman and Helpman (1991a), except that the last term on the right-hand side, $-\rho$, is not divided by λ. This change reflects the fact that we have removed the monopoly distortion effect. The monopoly distortion effect is higher, the higher is λ. This is intuitive, because the monopoly distortion effect reflects the mark up charged in the market for innovative good, which is proportional to λ.[20]

Proposition 3 *In the modified Grossman and Helpman model with no monopoly distortion effect, there cannot be over-investment in R&D in equilibrium.*

Proof. Since the economy grows at rate $g = \iota \log \lambda$, discounted utility can be directly calculated as

$$u = \frac{\log [L - \iota a]}{\rho} + \frac{\iota \log \lambda}{\rho^2} - \log L$$

To show that there is always over-investment in R&D in the market equilibrium, we must show that $\frac{du}{di} \mid_{i=i^*} > 0$ whenever $i^* > 0$, that is whenever $0 \leq a\rho \leq \frac{\lambda-1}{\lambda} L$. We have

$$\frac{du}{di} \mid_{i=i^*} \propto \left(1 - \frac{\lambda - 1}{\lambda}\right) L - a\rho \left(\frac{1}{\log \lambda} - 1\right).$$

The first term on the right-hand side is positive for any $\lambda \geq 1$. If $\lambda \geq e$, then the second term is always non-negative, so $\frac{du}{di} \mid_{i=i^*} > 0$. To complete the proof, assume that $1 \leq \lambda < e$. In this case, the second term monotonically decreases with $a\rho$. Therefore, we only need to show that $\frac{du}{di} \mid_{i=i^*}$ is non-negative at $a\rho = \frac{\lambda-1}{\lambda} L$. That is:

$$\frac{du}{di} \mid_{i=i^*, a\rho = \frac{\lambda-1}{\lambda}} = L\left(1 - \frac{\lambda - 1}{\lambda \log \lambda}\right) \geq 0$$

for $1 \leq \lambda < e$. To show that this inequality holds, consider first of all the case $\lambda = 1$. Taking the limit, we have:

$$\lim_{\lambda \to 1} \frac{\lambda - 1}{\lambda \log \lambda} = \frac{\lim_{\lambda \to 1} (1)}{\lim_{\lambda \to 1} (1 + \log \lambda)} = 1,$$

[20] The fact that the monopoly distortion effect increases with λ does not imply that over-investment is more likely when the size of innovations is large, as the countervailing effects, which tend to generate under-investment in R&D, are also increasing in the size of innovations. Thus, in the Grossman and Helpman (1991a) model, the monopoly distortion effect can prevail when the size of innovations is either very large or very small.

which means that $\left(1-\frac{\lambda-1}{\lambda\log\lambda}\right)$ tends to zero as λ goes to 1. To complete the proof it now suffices to show that $\left(1-\frac{\lambda-1}{\lambda\log\lambda}\right)$ monotonically increases in λ on $(1,e)$. Indeed, we have:

$$\frac{\partial}{\partial\lambda}\left(1-\frac{\lambda-1}{\lambda\log\lambda}\right)=\frac{1}{\lambda^2\left(\log\lambda\right)^2}\left(\lambda-1-\log\lambda\right).$$

This is always positive for $1\leq\lambda<e$. ■

An alternative way of eliminating the monopoly distortion effect is to assume that labour used in research is taxed at an appropriate rate. To demonstrate this alternative approach, we now return to Grossman and Helpman's original assumption that labour is the R&D input. Thus, $n(t)$ now denotes the amount of labour employed in research, and the parameter a is assumed to be constant, as in the original Grossman and Helpman model.

To determine the tax rate needed to eliminate the monopoly distortion effect, observe that when $p=\lambda w$, the marginal productivity of labour, in terms of the final good, is λw. To ensure that the labour cost perceived by research firms is equal to the marginal productivity of labour, one needs to assume that labour used in research is taxed at rate $\tau=\lambda-1$. The fiscal revenue $(\lambda-1)w(t)a\iota$ is paid back to market leaders as a lump-sum subsidy that adds to their profits.[21]

With this tax in place, the free-entry condition in patent races becomes

$$v(t)=\lambda w(t)a,$$

and the profits obtained by the incumbent in each industry are

$$\pi(t)=\frac{\lambda-1}{\lambda}y(t)+(\lambda-1)w(t)a\iota.$$

To solve the model, we use the free entry condition, the no-arbitrage condition, the Euler equation, and the labour market clearing condition

$$L=a\iota+\frac{y(t)}{\lambda w(t)},$$

where the first term on the right-hand side is labour used in research, and the second is labour used in the production of intermediate goods. Oner can immediately see that we get exactly the same market equilibrium condition as in the lab-equipment model. The same conclusion therefore holds.

[21] The alternative assumption that the fiscal revenue is returned to consumers as a lump-sum subsidy would further decrease the investment in R&D, reinforcing the under-investment result.

Appendix B

This Appendix shows that there can never be over-investment in research in the quality-ladder models of Barro and Sala-i-Martin (2004) and Acemoglu (2009). For brevity, we shall adopt the original notation of these authors and use directly their formulas.

B.1. *Barro and Sala-i-Martin (2004)*

The rate of growth of the economy in the market equilibrium is (equation (7.36) at p. 330):

$$\gamma^* = \frac{(g-1)(\frac{\Pi}{\xi} - \rho)}{1 + \theta(g-1)} = \frac{\frac{\Pi}{\xi} - \rho}{\theta + (g-1)^{-1}},$$

where q is the step size of innovations, α is income share of capital, $g \equiv q^{\frac{\alpha}{1-\alpha}}$, ρ is the rate of time preference, θ is the intertemporal elasticity of substitution, ξ is an R&D cost parameter, and

$$\Pi = LA^{\frac{1}{1-\alpha}} \left(\tfrac{1-\alpha}{\alpha}\right) \alpha^{\frac{2}{1-\alpha}} \quad \text{if } q \geq \tfrac{1}{\alpha} \text{ (drastic innovations)}$$
$$\Pi = (q-1)\, LA^{\frac{1}{1-\alpha}} q^{-\frac{1}{1-\alpha}} \alpha^{\frac{1}{1-\alpha}} \quad \text{if } q < \tfrac{1}{\alpha} \text{ (non-drastic innovations)}$$

is the innovator's profit flow.

Following Barro and Sala-i-Martin (2004), we compare the equilibrium rate of growth γ^* to the socially optimal one, which is (equation (7.59) at p. 341):

$$\widehat{\gamma} = \frac{\frac{(g-1)}{g}\frac{S}{\xi} - \rho}{\theta},$$

where

$$S = LA^{\frac{1}{1-\alpha}} \left(\frac{1-\alpha}{\alpha}\right) \alpha^{\frac{1}{1-\alpha}}.$$

To prove that $\widehat{\gamma} \geq \gamma^*$ it suffices to show that

$$(g-1)S \geq g\Pi$$

B.1.1. *Non-drastic innovations*

When innovations are non-drastic, i.e. $q < \frac{1}{\alpha}$, inequality $(g-1)S \geq g\Pi$ becomes:

$$(g-1)LA^{\frac{1}{1-\alpha}} \left(\frac{1-\alpha}{\alpha}\right) \alpha^{\frac{1}{1-\alpha}} \geq g\,(q-1)\, LA^{\frac{1}{1-\alpha}} q^{-\frac{1}{1-\alpha}} \alpha^{\frac{1}{1-\alpha}}$$

which reduces to

$$\frac{1-\alpha}{\alpha} \geq \frac{q-1}{q(q^{\frac{\alpha}{1-\alpha}} - 1)}$$

To prove that the inequality always holds, we first show that the right-hand side decreases in q for $q \in [1, \frac{1}{\alpha}]$. We have:

$$\frac{\partial}{\partial q}\left(\frac{q-1}{q(q^{\frac{\alpha}{1-\alpha}}-1)}\right) = \frac{1}{q^{\frac{2-3\alpha}{1-\alpha}}(1-\alpha)\left(q^{\frac{\alpha}{1-\alpha}}-1\right)^2}\left(-q\alpha + q^{\frac{\alpha}{\alpha-1}}\alpha - q^{\frac{\alpha}{\alpha-1}}+1\right).$$

Therefore,

$$\frac{\partial}{\partial q}\left(\frac{(q-1)}{q\left(q^{\frac{\alpha}{1-\alpha}}-1\right)}\right) \propto 1-\alpha q-(1-\alpha)\,q^{-\frac{\alpha}{1-\alpha}}.$$

Notice that the right-hand side is 0 at $q=1$ and decreases with q for any $q \in (1, \frac{1}{\alpha}]$. Therefore, $\frac{\partial}{\partial q}\left(\frac{(q-1)}{q\left(q^{\frac{\alpha}{1-\alpha}}-1\right)}\right) < 0$ for any $q \in (1, \frac{1}{\alpha}]$.

Next, notice that, using L'Hospital's rule:

$$\lim_{q\to 1}\frac{(q-1)}{q\left(q^{\frac{\alpha}{1-\alpha}}-1\right)} = \frac{\lim\limits_{q\to 1} 1}{\lim\limits_{q\to 1}\left(q^{\frac{\alpha}{1-\alpha}}-1+\frac{\alpha}{1-\alpha}q^{\frac{\alpha}{1-\alpha}}\right)} = \frac{1-\alpha}{\alpha}.$$

It follows that $\frac{1-\alpha}{\alpha} > \frac{q-1}{q(q^{\frac{\alpha}{1-\alpha}}-1)}$ whenever $1 < q < \frac{1}{\alpha}$, completing the proof of the over-investment result in the case of non-drastic innovations.

B.1.2. *Drastic innovations*

With drastic innovations and monopoly pricing, i.e. $q \geq \frac{1}{\alpha}$, inequality $(g-1)S \geq g\Pi$ becomes:

$$\frac{q^{\frac{\alpha}{1-\alpha}}-1}{q^{\frac{\alpha}{1-\alpha}}} > \alpha^{\frac{1}{1-\alpha}}.$$

Since the left-hand side is increasing in q, it suffices to prove that the above inequality holds for $q = \frac{1}{\alpha}$. This is equivalent to:

$$\alpha^{-\frac{\alpha}{1-\alpha}} - (1+\alpha) > 0$$

which is indeed always true for $\alpha \in (0,1)$.

B.2. *Acemoglu (2009)*

For brevity we focus on the case of drastic innovations (as Acemoglu himself does), but the same conclusions hold with non-drastic innovations. Acemoglu's specification of the final good technology implies that the condition for innovations to be drastic is $\lambda \geq \left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}}$. The market equilibrium and socially optimal growth rates are, respectively

$$g^* = \frac{\lambda\eta\beta L-\rho}{\theta+(\lambda-1)^{-1}}$$

and

$$\hat{g} = \frac{\eta(\lambda-1)(1-\beta)^{-\frac{1}{\beta}}\beta L - \rho}{\theta}.$$

In order to demonstrate the possibility of over-investment, Acemoglu uses the following numerical example:

$$\theta = 1, \beta = 0.9, \lambda = 1.3, \eta = 1, L = 1, \rho = 0.38.$$

With these parameter values, he correctly calculates $g^* = 0.18231$. However, the value of $\hat{g}$ reported in Acemoglu (2009, p.467) is not correct. The correct value is

$$\hat{g} = (1.3-1)(0.1)^{-\frac{1}{0.9}}(0.9) - 0.38 = 3.107\,2,$$

which is greater than g^*.

To prove that in fact g^* can never exceed $\hat{g}$, it suffices to show that

$$\lambda\eta\beta L < \eta(\lambda-1)(1-\beta)^{-\frac{1}{\beta}}\beta L.$$

The economic meaning of this condition is simply that the private value flow of an innovation is always lower than its social value flow. This inequality can be re-written as:

$$\frac{\lambda}{(\lambda-1)} < (1-\beta)^{-\frac{1}{\beta}}.$$

Since $\frac{\lambda}{\lambda-1}$ decreases with λ, if the inequality is satisfied for $\lambda = \left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}}$ (the lowest size of innovation such that innovations are drastic), it is also satisfied for $\lambda > \left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}}$. Then, all we need to show is that

$$\left(\left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}} - 1\right)(1-\beta)^{-\frac{1}{\beta}} - \left(\frac{1}{1-\beta}\right)^{\frac{1-\beta}{\beta}} > 0$$

for $\beta \in (0,1)$, which is indeed always true.

Appendix C

This Appendix shows that there cannot be over-investment in R&D in the Jones and Williams (2000) model once the R&D congestion effect has been removed. The R&D congestion effect can be eliminated simply by assuming that there are constant returns to R&D. Using Jones and Williams' notation, this is equivalent to setting the parameter λ equal to 1.

Setting $\lambda = 1$, the market equilibrium and social optimum are respectively given by:

$$s^* = \frac{\pi\frac{A}{Y}(1+\psi)g_A}{r - g_Y + (1+\psi)g_A}$$

$$\hat{s} = \frac{\sigma g_A}{r - g_Y + (1-\phi)g_A},$$

where s is the share of R&D in national income. The meaning of the symbols here is as follows: r is the interest rate, ρ is the elasticity of substitution, α is the labour share of income, ψ is the number of upgrades per new variety,

$$\pi = \frac{\eta - 1}{\eta}(1-\alpha)\frac{Y}{A}$$

is the innovator's profit, where η is the mark-up and $\frac{Y}{A}$ is the output/stock of knowledge ratio, which is constant in a steady state, ϕ is an intertemporal spillover parameter,

$$\sigma = \frac{1}{\rho} - (1-\alpha),$$

$$g_A = \frac{n}{1 - \phi - \frac{\sigma}{\alpha}}$$

where n is the rate of growth of labour supply, and

$$g_Y = \frac{\sigma}{\alpha} g_A + n.$$

To show that there cannot be over-investment in R&D, we prove that the numerator of s^* cannot exceed that of $\hat{s}$. It is obvious that the denominator of s^* is necessarily greater than that of $\hat{s}$, so this suffices to show that $s^* < \hat{s}$. Consider first the standard case where $\psi = 0$. In this case, the equilibrium mark-up η is given by

$$\eta = \frac{1}{\rho(1-\alpha)} = \frac{\sigma}{(1-\alpha)} + 1.$$

The numerator of s^* then becomes

$$\frac{(1-\alpha)}{\sigma + (1-\alpha)} \sigma g_A,$$

which is clearly lower than σg_A, i.e. the numerator of $\hat{s}$. Intuitively, with simple monopoly pricing the private value of an innovation must be lower than its social value.

Now consider the case with "upgrades." By bundling upgrades and new varieties, the innovator can spread monopolistic distortions over a greater number of goods. It follows from standard Ramsey pricing logic that he can extract more surplus, as in Burstein (1960). As noted by Jones and Williams (2000, p. 71), the greater is the number of upgrades, the more evenly the monopolistic distortion can be spread, and hence the greater is the surplus that can be extracted. However, the latest innovator must price the bundle in such a way that buying the bundle is preferable to buying only the upgraded varieties at a price equal to marginal cost. (Since the original inventors of those varieties are displaced by the last innovator, they must stand ready to provide those varieties at a price equal to marginal cost.) This places an upper bound on the surplus that the latest innovator can extract, which cannot exceed the social value of the innovation.

To show this, notice first of all that the upper bound is attained in the limit as $\psi \to \infty$, that is when the number of upgrades is arbitrarily large. Using Jones and Williams's formula for the mark-up with upgrades, namely

$$\eta = \left(\frac{1+\psi}{\psi} \right)^{\frac{\sigma}{1-\alpha}},$$

the numerator of s^* becomes

$$\frac{\eta - 1}{\eta}(1-\alpha)(1+\psi)g_A.$$

It is then easy to calculate

$$\lim_{\psi \to \infty} \frac{\eta - 1}{\eta}(1-\alpha)(1+\psi)g_A = \sigma g_A,$$

which proves that $s^* < \hat{s}$.