Basic Research and Prosperity: Sampling and Selection of Technological Possibilities and of Scientific Hypotheses as an Alternative Engine of Endogenous Growth

Cornelis A. van Bochove

CWTS Working Paper Series

Paper number	CWTS-WP-2012-003
Publication date	March 13, 2012
Number of pages	60
Email address corresponding author	cbochove@cwts.leidenuniv.nl
Address CWTS	Centre for Science and Technology Studies (CWTS) Leiden University P.O. Box 905 2300 AX Leiden The Netherlands www.cwts.leidenuniv.nl

Universiteit Leiden

CWTS

Centrum voor Wetenschaps- en Technologie Studies

Basic Research and Prosperity: Sampling and Selection of Technological Possibilities and of Scientific Hypotheses as an Alternative Engine of Endogenous Growth.

Cornelis A. van Bochove,

Centre for Science and Technology Studies, Leiden University

March 2012

Abstract

In endogenous growth theory technological progress is generated by accumulation of knowledge. This causes a number of problems, most importantly the scale dependence of technological progress. This paper develops a growth model where technological progress is based on sampling and selection from a Pareto distribution of technological possibilities. If sampling is purely random, slow growth is possible provided population increases, explaining the pre-industrial growth experience. If accumulation of experimental sampling mass is enough to boost the chance of success in further sampling, the explosive regime of endogenous growth is reproduced, explaining the double digit growth of transitional countries. The moderate growth of advanced economies is reproduced if basic research, modeled as a process of sampling and selection from a Pareto hypothesis distribution, is needed to improve the chance of experimental success. The rate of technological progress then mainly depends on a balance between the power of the technology function, which indicates how hard further innovation is, and the rate of learning to develop and test new hypotheses in basic research. A high rate of population growth somewhat increases the rate of technological progress, but there also is growth with a stable population. Directing basic research towards economic opportunities is detrimental to growth and may reduce the growth rate by as much as one half. The steady state is shown to be globally stable; in the steady state, the growth rate is independent of the research intensity, but the level of income depends on it. Given current OECD levels of R&D spending and saving, a one dollar increase of applied R&D spending will increase national income with 6-25 dollars and one dollar extra basic research by 20-100 dollars. These rates of return are ten and thirty times higher, respectively, than those on physical capital investment.

Keywords: Endogenous growth, basic research, science, technological progress, applied R&D, technology function, hypothesis function, innovation, learning, research intensity.

JEL classification: D24, O31, O33, O40

1. Introduction

Current levels of prosperity in the advanced countries are without historical precedent. It is uncontroversial that this is due to technological progress based on scientific discoveries. Yet the mainstream neoclassical theory of economic growth is almost silent on the role of basic research in economic growth. This theory is based on the premise that technological progress is generated by the accumulation of a stock of technological knowledge or 'ideas'; apart from a few very recent papers, the literature does not differentiate the stock of knowledge into basic knowledge, applied knowledge or practical knowledge. This lack of resolution of the theory is a major problem in science and technology policy: there is no theory to guide the allocation of budgets to these three purposes. As a consequence, budgetary policies with respect to this field as well as priority setting within them are matters of taste and fashion rather than of rational decision making. In the present section we first briefly review the existing growth literature, then introduce our alternative and finally summarize our findings.

The canonical endogenous growth literature

That technological progress, not capital investment, is the ultimate driver of growth was established more than half a century ago by Solow (1956). This means that to understand economic growth one needs to understand technological progress. The current standard theory of technological progress is endogenous growth theory, also known as AK theory, (neo-) Schumpeterian theory and evolutionary growth theory. This theory has roots that go back to the 1970's (cf. Paul Romer, 1994), found a canonical form in the 1990's with the work of Romer (1990), Grossman and Helpman (1991), Aghion and Howitt (1992, 1998), Charles Jones (1995a, b), Kortum (1997), Dinopoulos and Thomson (1998), Peretto (1998) and Young (1998), Segerstom (1998). At the end of the decade it was reviewed concisely and incisively by Charles Jones (1999) and won its proper place in macro-economic textbooks (David Romer, 2001).

The basic approach is to assume that there exists a stock of knowledge or ideas that determines the level of technology and that this stock of knowledge can be increased by knowledge production in which the stock of knowledge itself plays a central role. There are many variations on the theory, but its core results are quite simple. Let the production function be $Y = Af(K, L)$ where Y is national production of goods and services, K physical capital, L labor and A the stock of knowledge. The notation A goes back all the way to Solow's (1956) seminal paper, although he interpreted A not as the stock of knowledge but simply as a scalar that determines the level of factor productivity. Capital accumulates neo-classically, with fixed rate of saving; the stock of knowledge grows according to $\dot{A} = L_A A^{\phi}$. Here a dot indicates a derivative with respect to time, L_A is the labor force engaged in knowledge accumulation and ϕ is a constant, positive, parameter. The long term growth of the economy now sensitively depends on the value of ϕ. If $\phi = 1$, and the knowledge producing labor force is constant, there is constant rate of technological progress and hence of long term economic growth; this rate depends on the size of the research labor force. This implies that the rate of growth depends on the expenditure on knowledge accumulation, which is intuitively appealing, but also that there is a scale effect: the growth rate depends on the size of the economy since expenditure on knowledge creation depends on it. Therefore a small economy has a lower rate of growth than a large one. Similarly, if the knowledge labor force grows exponentially, so does the rate of technological progress, causing per capita income to grow explosively. This explosive growth is lost if ϕ is smaller than one. In this case, development of new knowledge becomes more difficult as the stock of knowledge grows or, in Jones's (1995) terminology, 'fishing out' makes new ideas more difficult to develop. Then, if the knowledge labor force grows at the same exponential rate n as the total labor force, the rate of technological progress converges to $n/(1-\phi)$ and, with f Cobb-Douglas, the rate of growth of per capita output tends to $\frac{n}{(1-\phi)(1-\alpha)}$, where α is the factor share of capital. This means that the scale dependence has not vanished: per capita income grows only if there is positive population growth and on the steady growth path the levels of knowledge and income depend on the size of the population. If $\phi > 1$, in contrast, we once more obtain explosive growth, even if there is no population growth. In this case new knowledge becomes easier to develop if the existing stock of knowledge is larger, the 'standing on shoulders' effect (Jones, 1995).

Following Jones (1995, 1999), the endogenous growth literature identifies the dependence on scale as an undesirable feature, though there are some exceptions (e.g. …). Dependence of the rate of technological progress on the rate of population growth or even the absolute size of the economy is hard to reconcile with the historical experience in the advanced countries in the past two centuries: a moderate rate of per capita growth, irrespective of the growth of the populations and the size of the economies. Explosive growth is not realistic in the long either. It is a good description of a transitional development phase (notably Japan until 1980, China, India and Brazil today) but this would mean that ϕ changes with the stage of development and thus would require a new model with an endogenous ϕ. Moreover, explosive growth in the transitional countries mostly occurred after populations were stabilized and were certainly no direct consequence of rapid population growth. To counter these objections a way has to be found to 'rescale' technological progress. In the canonical endogenous growth literature cited above, this is done by assuming that the scale effect on R&D and technological progress is offset by a factor that increases the difficulty of innovation if the economy is larger. In particular, it is assumed that there is an increase in product diversity. If product diversity increases at exactly the same rate as the population and if for every product knowledge accumulation grows according to the equation with $\phi = 1$, the explosive pressure of labor force growth or, in an alternative specification of the model, of technology driven income growth is precisely countered and the rate of technological progress once more can be constant, independent of n and dependent on the budget allocated to knowledge accumulation.

The knowledge-accumulation-based models have a number of undesirable properties that hamper their effectiveness as underpinning of research and innovation policy. They depend sensitively on precise values of parameters that are hard to measure. Stable growth depends not merely on the precise value of ϕ but also on the precise balancing of labor force growth by increasing product diversity. The ad hoc nature of this assumption diminishes the explanatory power of the theory; arguably, steady state growth is assumed rather than explained. Conceptually, the assumption that product diversity depends on size as measured by population is not easy to reconcile with actual product diversity in European countries or Australia compared to the much larger US. Even if the world as a whole is considered, the models paradoxically imply that explosive growth could easily be obtained by limiting product variety.

Extensions of the literature

An alternative mechanism to prevent explosive growth was developed recently by Ben Jones (2010). He assumes that the time required to master the stock of knowledge in a field increases with the maturity of the field (his terminology is an increasing 'burden of knowledge'), implying that individuals need to invest more time to become productive in innovation. Individuals choose fields, taking account of the burden of knowledge in a field, to maximize their lifetime chance of success, leading to a balance with the effect of increased availability of innovation resources induced by economic growth. A drawback of Jones's model is that his steady state growth rate (of which the stability is not established) depends on the existence of population growth. Without it, his growth rate is zero. Thus there still is a scale effect. More important, however, is that Jones's this line of reasoning is difficult to reconcile with the cases where very major recent innovations have been made by very young people (e.g. the PC, internet search engines, social media) and with the fact that a large part of international academic articles is authored by junior researchers (PhD students and post docs), particularly in the sciences. It should be recognized that mastering existing knowledge, even if it is recent, is several orders of magnitude easier than discovery of new knowledge; this means that it is hard to see how the need to possess existing knowledge could explain a slowdown of discovery. In fact, it would seem that any realistic theory of technological progress should explain why transmission and use of existing knowledge is so much easier than development of new knowledge. Jones's empirical arguments (notably that the age at which Nobel winning work and breakthrough patents are achieved has increased in the course of last century may very well be due to another factor entirely. R&D probably requires more technological resources today than a century ago and it has become embedded in much larger organizations. Therefore, command over resources and over the research agenda has now, just as in non-R&D sectors, become more like a management position achieved not much before the age of forty, rather than a matter of talent and creativity. Thus breakthrough work is being done at a later age not because much more knowledge needs to be mastered but because research freedom and availability of the needed technological resources comes at a much higher age than before.

Since the canonical literature has been established a growing number of papers has been published with extensions, using the original basic mechanism for technological progress. An early review of some new developments was given by Dinopoulos and Şener (2007) but since then more has appeared. The presence of product diversity in endogenous growth theory, modeled as an easy to analyze continuum, has spawned a large number of papers harnessing that feature to describe phenomena other than macro-economic growth. It would be well worth the effort to write a full scale review of the literature focusing on the use of this mechanism, but this would be far beyond the scope of the present paper. Instead we just mention some important threads, and give some of the many references by way of illustration, to give an inkling of the scope of this literature. One important knowledge-related thread in the literature is to consider how patenting systems prevent spillovers from making innovation work at the product level; this is analyzed by postulating that each product is produced by a monopolist in need of protection of the fruits of his R and D expenditure. Other threads link R&D to the business cycle and financial markets (*e.g.* Dosi *et al*., 2010, 2012), consider the funding role of government, and extent the model to open economies. From the point of view of science and research policy the most important extensions are recent attempts to differentiate between basic and applied R&D: Cozzi and Galli (2009), Gersbach *et al*. (2008, 2009, 2010). Most elaborate is Gersbach *et al*. (2009). Their model is a variation of the original constant labor force endogenous growth model. The extension is that the production of knowledge is refined. A is a function of the available number of intermediate goods. To increase the latter requires first of all basic research results and on top of that applied research to turn them into actual intermediate products. Naturally, in this model growth ceases without basic research. But, though that case is not developed, growth should also become explosive if basic research and labor force growth are combined. The scale problem remains.

Empirically, much work has been done (summarized by Donselaer, 2011), but the results are not uniform. Basically, differences in the rate of technological progress measured as the residual of the output growth less the sum of the growth contributions of non-knowledge production factors (Solow's 1957 approach), cannot strongly support a particular one of the endogenous growth models. The evidence clarifies that R&D, population growth, human capital and interactions between countries have been are important factors in the advanced countries, but that the total (applied plus basic) research intensity does not play the dominant role predicted by endogenous growth theory (Madsen *et al*., 2010). In contrast, in case of the Asian transitional double digit growth countries, there is good evidence (Madsen, 2010; Ang and Madsen, 2011) that the research intensity is as important as endogenous growth theory indicates and is related directly to the growth rate.

Our approach: R&D is not knowledge accumulation but sampling and selection

The conceptual basis of the existing models is shaky in several respects. Whereas it is not impossible to measure the national stock of physical capital, it is hard to see how a similar measurement could be made of the national or global stock of knowledge (Steedman, 2003). New knowledge may replace existing knowledge, encompass it, enhance it, cause it to be forgotten, build on it, and so on. Therefore there is no simple relation between the stock of knowledge and R&D. The latter is not simply the increment, gross or net, of the stock of knowledge. Similarly, Griliches (2000) in the book that summed-up his life's work in the field, noted that R&D 'capital' does not simply depreciate as a consequence of the mere passage of time or of mechanical wear and tear, but can become obsolete simply because of new work. It is no coincidence that Jones (1995, 1999) did not even use the word stock of knowledge but preferred 'stock of ideas'. The problem to define a stock of knowledge is exacerbated if one wishes, as makers of science and innovation policy must in their funding decisions, to distinguish between various types of knowledge or ideas. How does one add up the knowledge created in basic discoveries and knowledge created to achieve a specific application?

This conceptual problem reflects the fact that the concepts of accumulation and saving that have been developed to describe capital formation are less relevant in the field of science and technology; mutatis mutandis the same is true of the concepts of product diversity and intermediate goods that have become an important tool in the literature. In both cases, models and terminology that are natural in the fields for which they were originally designed, are unnatural straightjackets in the analysis of R&D. This includes words like accumulation, competition, production, intermediate products, optimization; an outstanding example of the mismatch between the field and the analytical apparatus is the use of the term 'spillover effects' in case of basic research: it suggests that having their work picked up by others is mainly a nuisance to basic research workers! Much more relevant and natural in the field of R&D are concepts such as discovery, trial and error, serendipity, transmission, speed of learning, laws of nature, theories of everything, fractionalization of knowledge fields, and so on. Therefore we introduce a theory that starts out from these concepts and use them to obtain an alternative neoclassical model of the generation of technological progress and economic growth.

To achieve this we do away with the assumption that technological progress is produced or even *created* through accumulative processes. Instead we recognize that it is the result of the *discovery* of technological possibilities, possibilities that exist irrespective of whether human beings discover them or not. We emphasize that discovery is the result of a trial and error process where samples are drawn from all technological possibilities and the best results are retained. Knowledge then is either of two things. First, a record of the results of previous experiments; second, information on how to select a sample in such a way that the chance of success is increased. Consequently, two things are central in our approach: the inherent distribution of the technological possibilities and the sample selection process. With respect to distribution of the technological possibilities, it is clear that there are many different technologies that can generate a low level of productivity, but far less that result in a high productivity. A natural way to formalize this idea is to postulate the existence of technology density function that indicates what the relative number of technologies is that can generate a given level of productivity. We assume that his function has the shape of an inverse power law aka Pareto distribution. For growth analysis this distribution is particularly attractive because its shape is the same everywhere. The power k of the technology function indicates how fast the density falls with increasing productivity and thus how fast innovation becomes more difficult. Given the technology function, technological progress can be described as sampling and selection regimes. Sampling corresponds with experimental development (applied R and D) of new technologies. The various basic models encountered in the endogenous growth literature for different values of ϕ are now obtained as a consequence of differences between sampling and selection regimes, irrespective of parameter values. These regimes are introduced in sections 2 and 3 and their growth implications are analyzed in sections 4 - 7. We briefly summarize the core features and results for these regimes here.

- The simplest possible case occurs when there is no permanent sampling process but the first working technology found is used and no information is exchanged between regional communities; presumably this was the situation in the gathering and hunting societies that preceded the first agricultural economies. In this case the level of technology is simply the expected mean of the technology distribution and we have the original zero technological progress Solow model.
- If experiments are unconstrained by prior information, but a fixed part of income is spent on them and the experimental result with highest productivity is retained, there is a stable but slow growth path provided population growth is non-zero (cf. section 4); this is similar to the case $\phi < 1$ in the endogenous growth literature and to Jones's (2010) model. It can be identified with the growth process in pre-industrial agricultural societies, where some experimentation went on but no systematic development of the experimental process itself.

- In the next regime the information from the sampling process itself is sufficient to constrain the sampling process to exclude technologies with a productivity below the level where a critical experimental (sample) density has been achieved. This regime leads to the explosive growth corresponding to $\phi > 1$ (cf. section 5), even if population does not grow. This corresponds to the double digit growth process of transitional countries and explains the results of Madsen and others cited above on the relation between the research intensity and growth in Asian transitional economies.
- The final regime introduces basic research. After the onset of the industrial revolution, the accumulation of experimental mass is no longer enough in the advanced countries to constrain further experiments to potentially successful ones: it is necessary to discover the underlying laws of nature that constrain the technological possibilities. Unlike the endogenous growth literature implies, science's role is not to open up technological possibilities, but to guide technological R&D away from unsuccessful avenues and thus to increase its effectiveness. The classic example is that no amount of unsuccessful experimentation could deter the alchemists from trying to make gold; the discovery of the basic laws of chemistry within the periodical system was needed to do that. That this is no mean feat is illustrated by the fact that Isaac Newton devoted more years to alchemistic experiments of which we now know they were pointless, than to his scientific discoveries on light and gravity that helped shape our world.

  Basic research amounts to developing and testing of hypotheses until one is found that provides a sufficient explanation of the data, in this case the experimental results leading to a given level of productivity. We describe this process with a concept similar to the technology density function (cf. section 3): the hypothesis density function. Defining a research field as the development and testing of hypotheses to explain a given level of productivity (implying that there is an ever expanding continuum of research fields), a critical density of tested hypotheses has to be achieved. The speed with which this happens depends on the number of workers in the field and the fall of the time needed per hypothesis in that field due to learning. This leads to a stable steady state rate of growth of which the value depends on the balance between the speed of learning in new basic research fields and the increasing difficulty of innovation expressed in the power of the technology function (the steady state growth rate is given in section 6, its stability is shown in section 7). The growth rate also depends on the rate of growth of the labor force, but does not vanish if the latter is zero. This case is a bit like the case in the endogenous growth literature where $\phi = 1$ and product differentiation balances the tendency to explosive growth, but unlike the literature it also generates steady state growth without population growth.

The regime with basic research permits an analysis of some important policy issues. One is whether or not to target basic research. It has become fashionable in many countries to implement measures to direct basic research at specific business sectors or application fields that are deemed to be particularly promising economically. This amounts to a portioning of the basic research labor force according to fields. The alternative approach is to leave basic research free, implying that there is a kind of joint production process in all fields. It turns out that for a plausible range of parameter values targeted basic research leads to a much lower steady state growth rate than free basic research; a ballpark figure is that growth rate is halved by targeting (Cf. section 6).

Another analysis that can be done is the effect of research spending on the level of income, or, more appropriately, on national consumption (Cf. section 8). The effect of course depends on both the (unknown) parameter values and on the values of the rates of spending on applied R and D and basic research themselves. Surprisingly, if these rates are small, the effect is completely dominated by the rates and virtually parameter free. For realistic cases, applied R and D will have a return of something like ten times that of spending on physical capital and basic research will have return that is 3 to 4 times as high as that on applied R and D. At current rates of spending of basic research of half a percent of GDP or less, one additional Dollar or Euro spent on basic research will raise national consumption by 50 to 100 dollars or Euros.

Finally, the importance of basic research and its processes have science policy consequences. The dependence of the long term growth rate on the speed of learning in basic research implies that enhancing this rate of learning is important for economic growth. This could be achieved by measures to speed up the dissemination of research results, such as immediate open access and electronic accessibility of all publications (for an estimate of the effects of free access cf. Houghton and Sheehan, 2009); full accessibility of data; complete publication of negative results; systematic reviews that are easy to understand for non-specialists, including outreach activities explaining results to a more general public; frequent contact between researchers in different countries, institutes and fields; reduction of administrative burdens; loss of research experience by people dropping out of the research system because of lack of career opportunities, and so on. None of these measures are new, many of them have been and are advocated by learned societies, but our analysis implies that they should be regarded as not just nice to do, but as central to effective national and international innovation policies, perhaps instead of part of the subsidies and tax breaks that many countries now spend their innovation budgets on.

The briefest possible summary of our results is that President Obama was right when in his 2012 State of the Union Address he declared, somewhat at odds with much of established theory and with current policies in most of the rest of the world, 'Innovation also demands basic research.' By our results, basic research is the dominant engine of prosperity in advanced nations.

2. The technology function and the sources of technological progress

In this section we introduce the concept of a technology function as a representation of the technological possibilities underlying past, present and future actual technologies. Furthermore we describe the processes of trial and error by which the actual technologies are selected from the technology distribution.

Technology function

Consider the standard aggregate production function where output or income Y is the total factor productivity A multiplied by a linear homogenous function y of capital K and labor L:

(1) $$Y = Ay(K, L)$$

For any individual good or service, the total factor productivity may differ considerably between producers, localities, countries and dates. This can be due to both major and minor differences in technology and production processes. Very high and very low productivities may co-exist at the same date, and productivities may have been higher in some localities thousands of years ago than they are in other localities today (e.g. due to conditions of climate, soil, mineral and natural resources, presence of waterways, etc.). This variation implies that even at the micro level total factor productivities can be considered as a sample from a random distribution. *Eo ipso*, the aggregate total factor productivity can be considered as the average of a large number of samples from an underlying distribution. Houthakker (1953) already proved that a macro Cobb-Douglas function results from Pareto-distributed fixed coefficients technology at the individual level. We take this one step further and assume that the aggregate total factor productivity A itself is the average of a sample from an underlying distribution. This assumption implies that A depends on two things: the underlying distribution and the sampling process.

The underlying distribution is formalized by a probability density function $f(x)$, where x is a value of total factor productivity. In economic terms, this function may be viewed as measure of the number of different technologies capable of generating total factor productivity x. In other words, the higher the value of $f(x)$, the easier it is to achieve a factor productivity value of x. We will refer to this density function as the *technology function* and use the term *technology point* x to indicate the technologies underlying a total factor productivity value x. We make three general assumptions about its properties. The first is a positive minimum value x_{min} of total factor productivity; below the minimum the density is zero, ruling out negative output: $f(x) = 0$ for $x < x_{min}$. The second assumption is that there is no maximum total factor productivity and therefore no absolute limits to technological progress and economic growth: $f(x) > 0$ for $x \geq x_{min}$. In its absolute form this assumption requires that no 'theory of everything' exists or, alternatively, that such theory allows infinite technological possibilities. This is the 'endless frontier' assumption of Vannevar Bush (1945) on which US science and technology policy has been based ever since he phrased it. Of course it is impossible to verify this assumption. However, it is a simple summary and a straightforward extrapolation of the historical experience until today. Nevertheless we temper it with a third assumption, namely that the technology function decreases as a function of the value of total factor productivity: $\frac{df(x)}{dx} < 0$ for $x \geq x_{min}$. This implies that as higher and higher levels of factor productivity are achieved, further technological progress, though possible in principle, requires more and more effort. Our final assumption is that the mathematical expectation of $f(x)$ is finite. If not, even a completely random trial and error process would immediately lead to infinite total factor productivity, in evident contradiction to historical reality.

These assumptions are satisfied by various distributions, including the upper half of the Normal distribution. We use the Pareto distribution. It is not only quite tractable but also embodies the power laws that usually pop-up in bibliometrics and scientometrics, most famously in Lotka's (1926) law, which states that the number of papers per author is inversely related to the square of the number of authors, implying a Pareto distribution with a power of 2. The impact factors of articles and authors, the total number of publications of authors, as well as many other phenomena in knowledge fields have long tail distributions and appear to be 'scale free' in the sense that they look the same everywhere; examples are given by *e.g.* Poiter (1981), Ivancheva (2001), Glänzel (2010), Li (2002). The Pareto distribution, perhaps generalized to account for another shape before the tail, is the natural form to describe this. The long tail phenomenon in its turn may be generated by underlying processes such as interactions in scale free networks (). In our case only the tail is important and we can use the original Pareto form. Thus we have as probability density function (technology function) $f(x)$, and cumulative density function $F(x)$:

(2) $$f(x) = \frac{kx_{min}^k}{x^{k+1}} \ (x \geq x_{min} \geq 0, k > 1); \ F(x) = 1 - (\frac{x_{min}}{x})^k$$

Here k is the power of the distribution; the higher k, the steeper the distribution decreases as a function of the total factor productivity: $d \ln f(x)/d \ln x = -(k+1)$. The restriction $k > 1$ assures existence of the mean.

Pure trial and error

We now turn to the sampling process that determines the actual total factor productivity. First, as a thought experiment, consider the purely static case. An example would be a large number of pre-agricultural societies with limited mutual communication, little opportunity to experiment with alternative means of production and almost no capacity for storing and transferring information about the results of experiments. Random differences would exist between the total factor productivities of societies but no systematic growth. On average the total factor productivity is simply the mean of the technology distribution:

(3) $$A = E(x) = \frac{kx_{min}}{k-1}$$ (Static case, pre-agricultural societies)

As soon as there is some communication between societies and some permanent capacity to experiment because a small part of income is set aside for experimentation, the expected result is no longer the mean of the distribution but the expected maximum productivity within the total cumulative sample. A newly tried out technology replaces the current one as soon as soon as it yields a higher total factor productivity. This trial and error process will, for economies as a whole and over longer periods of time, lead to an increase of the total factor productivity. The process is a plausible source of progress in agricultural societies before the advent of systematic research and development in the seventeenth century. Formally, the cumulative use of a part income to trial and error amounts to a continuous process of sampling. This, in turn, yields a growing cumulative sample which we will refer to as the experimental mass. The actual factor productivity can now be considered as the expected maximum of the cumulative sample. The maximum should be an increasing function of the experimental mass D. That this is indeed the case is shown in the appendix. The expected maximum in a sample of discrete size m from a Pareto distribution is:

(4a) $E(\hat{X}_m) = x_{min} \frac{m! k^m}{\prod_{j=1}^{m}(kj-1)}$

For large sample values we can approximate the discrete sample size m by the continuous experimental mass D and obtain:

(4b) $E(\hat{X}_m) \therefore D^{1/k}$ (Pure trial and error, pre-scientific agricultural societies)

Hence in a pure trial and error process the total factor productivity $A = E(\hat{X}_m)$ has a constant elasticity with respect to the (cumulative) experimental mass, the elasticity being the inverse of the power of the technology function. Clearly, if the power is high and the function as a steep slope, increasing the experimental mass has only a small effect on productivity. The assumption that D is continuous assumes that partial experiments are possible; since the technology function represents a national aggregate of many individual goods and services, each of which may have different technologies, this assumption is trivial. The economic growth of economies with pure trial and error technological progress is analyzed in section 3.

Knowledge-based trial and error

In a pure trial and error process the trials are a sample from the total distribution and there is no way to exclude the part of the distribution below the current total factor productivity from the sampling process. As a consequence, an ever larger part of the experimental technologies will turn out to have a lower productivity than current technology. In contrast, modern research and development does not try-out technologies at random, but makes use of accumulated knowledge and new research to obtain prior information to direct the experimental process towards the technologies that are most likely to lead to an increase in factor productivity. A natural way to stylize this is to stipulate that the sampling process eliminates a priori all technologies with productivities below a certain point, which we shall call the technology threshold. Then the expected total factor productivity is the conditional mean above the threshold. It is the ordinary conditional mean and not the expected sample maximum because in this case we want to emphasize knowledge driven improvement, not uninformed trial and error. In the case of the Pareto distribution, if the threshold is v $(v \geq x_{min})$ the conditional distribution and mean are:

(5) $f(x|x \geq v) = \frac{f(x)}{1-F(v)} = \frac{kv^k}{x^{k+1}}$ $\qquad$ $A = E(x|x \geq v) = \frac{kv}{k-1}$

Clearly (cf.4) the conditional distribution has precisely the same form as the original one, apart from being shifted to the right. This property of the Pareto technology distribution makes it ideal for the analysis of economic growth and the natural partner of the Cobb-Douglas production function.

Relation between applied R and D and basic research.

The specification of the impact of research and development on total factor productivity is now equivalent to specifying the mechanisms by which R and D affect the threshold v in equation (6). To derive a realistic specification of those mechanisms we need to consider the nature of various types of R and D in some detail. The most common distinction is that between basic research and applied R and D. In a nutshell, this is the distinction between discovering the laws of nature and harnessing them for practical purposes. The more elaborate standard definition is given in the OECD (2002, p.30) Frascati manual for R and D statistics. Basic research is defined as: 'experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts, without any particular application or use in view.' Applied research is defined similarly, but in this case 'directed primarily towards a specific practical aim or objective', and experimental development is defined as 'systematic work drawing on existing knowledge gained from research and/or experience which is directed to producing …' These categories of knowledge production are linked and need each other. To illustrate this we briefly consider two examples. The first is the classic and extremely well documented case of electromagnetism. Faraday's experimental work over several decades (ca 1810 - 1850) established the observational basis of electromagnetism. Maxwell (1865) then turned this into his unified electromagnetic field theory which subsequently facilitated the technological developments in the last part of the nineteenth century of Edison, Bell, Marconi, and others. Thus there was a sequence experiments → theory → applied R and D. The start of this sequence, however, was based in technology: many of Faraday's experiments were only possible because of newly available equipment and tools. Even so, he sometimes needed to do lot of development work in adjacent fields, e. g. in chemistry to develop the required materials. Even so, one of his experiments turned out to be impossible until new technological developments allowed Zeeman to do it in 1897.

The second example is drug development. Until fairly recently, drug development was based exclusively on applied R and D or even on trial and error: large numbers of substances were tried out as treatment of particular diseases in test tubes, in test animals and finally in clinical trials of increasing size and sophistication. This development was mostly done without basic knowledge of how healthy organisms function at the cellular and molecular level and not guided by knowledge of the mechanisms by which a disease disturbs the healthy organisms. In fact, in many cases of successful drugs, the knowledge of how and why they work was developed after they had proven to be successful, and then only partially. As a consequence, the selection of new substances to test could at best be based on hunches and analogies only. This has led to ever increasing cost of drug development, if only because without solid basic knowledge the only way to prevent lethal side-effects is intensive large scale testing. At present, no more than twenty new drugs are admitted for use every year, at a cost of 50 billion US$. Only in last decade and half has the technology of molecular biology developed so strongly that it is has become conceivable to obtain complete knowledge of how a disease works before designing a drug. This development is very much technology driven: sequencing techniques and other bio-chip technologies have made it possible to do lab tests that used to take years and were extremely expensive in a matter of days or even hours at negligible cost. These new techniques, based on advances in information technology and nanotechnology, are now causing such a deluge of data that the search for patterns and the development of models to understand the data is becoming a daunting task. Thus the technological advances are now gradually providing the basis for the advances in basic knowledge that might eventually lead to more rapid and cheaper drug development.

These two examples demonstrate three points:

- Applied R and D without sufficient basic knowledge will eventually become unproductive.
- Basic research requires an adequate level of technology and experimentation.
- These linkages are not simultaneous but sequential and require a lot of time.

Each of these properties are essential in the context of growth model: the model needs to take account of the fact that basic research in a field cannot start until a certain level of technology has been reached and a sufficient amount of experimentation has been done; applied R and D cannot be considered effective unless sufficient basic knowledge has been obtained; and the time lags need to be taken into account.

Critical density of applied R and D

It is almost trivial that the technology threshold v can only be increased after a sufficient amount of applied of R and D has been done at the corresponding the technology point. As before, we describe experimental work as the taking of random samples (experiments) from the underlying technology distribution; then we may formally define the *experimental density* $\delta(x,t)$ at a technology point x and time t as the number of experiments that led to precisely the value x of the total factor productivity. At any time, this density depends on the historical paths of both the experimental mass and the threshold. We again assume the experimental mass to be continuous and, moreover, sufficiently large for the difference between the expected density and the actual density to be negligible. Then at t, the change in the experimental density due to a marginal increase in the experimental mass and the density itself are:

(6) $$\frac{\partial\delta(x,t)}{\partial D(t)} = f(x)|x) \geq v(t)) = \frac{f(x)}{1-F(v(t))} \qquad \delta(x,t) = \int_0^t \frac{f(x)}{1-F(v(z))}\frac{\partial D(z)}{\partial z}dz$$

The first condition for an increase of the threshold, a sufficient amount of experimentation, can now be formulated more precisely: a necessary condition for the threshold v to increase is that the (expected) experimental density at $x = v$ is at least equal to a critical value δ_c. We will refer to this value as the critical density. We take δ_c to be constant and independent of x: differences in how difficult it is to develop new knowledge are already taken into account by the technology function**.** Define $x_c(t)$ as the 'critical point' at time t: the factor productivity value at which the density is precisely critical. In the appendix we show that the time derivative (time derivatives are denoted with dots) of the critical point is (all variables refer to time t):

(7) $$\dot{x}_c = \frac{k}{k+1}\frac{1}{\delta_c}\left(\frac{v}{x_c}\right)^k\dot{D}$$

As noted, the variable $D$is the mass of experiments. The production of experiments depends not just on labor, but also on capital and other material resources. In fact, because of the experimental and inductive, data driven, nature of applied R and D it is reasonable to assume that the experimental mass is produced with the same production function as output of ordinary goods and services. Therefore we can simply write

(8) $$\dot{D} = s_D Y,$$

Where Y is national income (or total output) and $s_D$the fraction of income spend on applied R and D; since we will use a neoclassical production function, this formulation implies that substitution is allowed between capital and labor in applied R and D.

3. Basic research

Production of basic research

The second condition for the technology threshold to increase beyond $x = v$, is that a sufficient amount of basic research results pertaining to the technology point x has been obtained. To obtain an appropriate measure of a 'sufficient amount' and how it is produced, we need to consider a number of aspects of the process of basic research. The first aspect is that of research fields. Each technology point represents an amalgam of production technologies of many different goods and services. Usually, increases in productivity do not occur to the same degree in all goods and services, but are concentrated in a number of industries, goods and services. At another technology point other industries, goods and services attain higher productivity gains. Therefore, at each point of the technology function, the required basic research results are based on a new mix of research fields. In fact, a continuous process of branching or fractionalization is exactly what is found empirically in scientometrics, cf. Van Raan (2000). This combination of shifts in the relative speeds of technological progress in industries, fractionalization of science, and intricate two-way relations between basic research and applied R and D, implies that each technology point x is associated with a unique mix of basic research fields, which we simply dub 'field x'. In view of the embedment of basic research in particular technology and experimental equipment, basic research in field x can only start once critical experimental mass has been achieved at technology point x.

The second aspect of basic research that needs consideration is the nature of this research. In textbooks, the exposition usually suggests an orderly progression of knowledge development where new contributions build on older work in a natural way, quite predictable by hindsight. Only very infrequently are there surprises, flashes of inspiration, caused by persons of extra-ordinary genius. This view of science was made canonical by Kuhn who implied that apart from the occasional paradigmatic change, the overwhelming majority of scientific work is more or less routine, straightforward work. The unpredictable part of research, the paradigmatic changes, were made even more out of the ordinary by stipulating that only young scientists made these changes and they are usually only accepted by the profession after the older generation has cleared the field.

It is however a cliché that basic knowledge in the making is something quite different from textbook knowledge. Actual research is much messier. Many things fail, successful things come as a surprise. Rather than a combination of normal accumulative work and very infrequent upheavals, it is continuous alternation of routine progress and upheavals of all sizes, staying with an approach for some time and hitting on and making paradigmatic changes of all kinds of magnitude. It is an amalgam of inspiration, groping in the dark, tenacious hard work, learning, and serendipity. There is not a dichotomy of a very small number of very major discoveries and very large number of routine stuff, but a power law, or at least a distribution with a long tail, where the pattern of new directions and routine continuation is repeated at every level in the same way. This is true per field, for the distribution of discoveries over individuals, and even for the distribution of discoveries within the work of individuals.

This process cannot be dealt with by a simple accumulation model. Instead, we again have to turn to a model of trial and error under uncertainty, with gradual learning. The units of this trial and error process are hypotheses about 'the underlying nature of phenomena end observable facts.' The process of basic research can be described as the development and elaboration of hypotheses, the analysis of their internal consistency, foundations, consequences and relation to other hypotheses, the testing of their compatibility with stylized knowledge of phenomena and detailed data, and their successfulness in prediction and in generating applications. This description can be summarized as 'developing and testing hypotheses'. To make this description amenable to quantitative analysis, we have to define a unit of measurement of hypotheses. This is quite straightforward in the context of the technology function model: all potential hypotheses in field x can be ordered according to the value h of the total factor productivity they generate. Here 'generate' means that they allow for raising the technology threshold to h. Having established this definition, we can model the trial and error process in a way that is similar to what we already did in case of technology: define a distribution function, then a selection process and finally derive the expected value of h. The probability density function for basic research hypotheses in field x, or 'basic knowledge function', is $g_x(h)$. Of course, this function could be any shape. However, in view of the preponderance of power laws in scientometrics, it is again natural to use a Pareto function. In principle, both the power and the minimum of this function could be field dependent. However, to avoid unnecessary extra parameters we assume a power that is the same for every field and, moreover, identical to the one of the technology function k. We also assume that the minimum is the same in every field. This makes sense if the complexity of the development of new technology derives from the complexity of the basic knowledge needed to underpin the technology. We can now write the basic knowledge function for field x:

(9) $$g_x(h) = \frac{kh_{min}^k}{h^{k+1}}$$

To complete the description of the process of basic research, we need to specify how the selection of hypotheses works. There are two major differences with applied research: with respect to the factors of production and with respect to learning. In case of basic research the critical production factor is labor. Though instruments have to be built in observational and experimental basic research, the restrictive factor usually is labor. Consider the development of molecular biology. The rapid development of DNA-chip technology has greatly increased the research possibilities, but this technology does not replace researchers, it allows them to do things that were not possible before. The critical role of labor is even stronger in theoretical work: Einstein's ten year hike from special to general relativity could not have been appreciably shortened by providing him with many more material resources (cf. Isacson, 2007). Consequently, a Harrod-Domar production function is natural for basic research; of course in this model material resources can also be the limiting factor, but in our case we can ignore this possibility since we already have assumed that critical experimental mass has to be achieved before basic research can start in a field.

The second major difference is learning. Later work in a field is easier because new analytical tools have been developed, unsuccessful approaches identified and unfamiliar concepts internalized. Also, earlier work is an input for later work and makes this easier. This learning process is reinforced in cases where growth in the volume of basic research in a field creates a market for technological tools which, once available, boost productivity. At present, this process is particularly strong in biomedical research, but it is also evident in *e.g.* ICT, nano-science, software production. In terms of hypothesis selection, the essence of learning is that the experience with the development and testing of hypotheses helps in selecting further hypotheses that have a higher chance of being successful than those in the original distribution. Thus, just as in applied work, there is a threshold that is raised by the research process. In basic research, this process is in one sense stronger than in case of applied research. Any researcher who succeeds in developing a successful hypothesis can immediately build on it in the selection of further work. However, in basic research there also is a restriction with respect to learning: it is an individual rather than collective process. Consider a thought experiment: suppose that not just Einstein would have worked on the general theory of relativity, but 1000 Einsteins who communicated only through the literature. Clearly, this would not have sped up the process of discovery much. Even if they would have cooperated quite closely, the gains would probably have been minor, because each of them would have had to try much the same approaches at the same time, each making comparable errors and exploring the same blind alleys before being convinced that they were errors and blind alleys. Isacson's (2007) description of the situation in 1915, ten years after the special theory of relativity, is quite instructive. Einstein had lectured on his approach and this had caused the greatest mathematician of the time, Hilbert, to work on the problem. Though this competition spurred Einstein on, it did not help him intellectually in any way at all, the solution he found was his and his alone. Of course this example cannot be completely generalized to all theoretical work; nor is all basic work theoretical. Yet it seems safe to say that in basic research, learning processes are sped up only a little by increases in the total mass of work done in the field but mostly depend on the cumulative experience of the standard worker in the field. This is not to say that the number of workers has no influence on output: given the state of learning, the number of workers determines the sample mass that is achieved in a field and therefore research output as long as that is defined as the maximum of the sample hypotheses.

We will now formalize these ideas in a model of basic research production. This consists of two parts: a learning equation and an output equation. To obtain the learning equation, let the average worker draw a series of successive samples from the hypothesis distribution and let the threshold in each sample be a function of expected results of the last sample. More precisely, let each successive sample require the same number of working hours ε and let the total number of hours worked by a standard worker in field x since work started be τ_x. Suppose each sample draw causes an increase in the threshold that reduces the remaining mass of the distribution with a fraction σ. Then we obtain from (2) for the relation between successive thresholds: $v(\tau_x) = v(\tau_x - \varepsilon)(1-\sigma)^{-\frac{1}{k}} = v(\tau_x - i\varepsilon)(1-\sigma)^{-\frac{i}{k}}$. If τ_x is a multiple of ε, setting $i = \tau_x/\varepsilon$ and noting $v(0) = h_{min}$ we obtain:

(10) $$v(\tau_x) = h_{min} e^{\frac{\theta}{k}\tau_x}$$

Here $\theta = -\frac{1}{\varepsilon}\ln(1-\sigma)$. Consequently, the threshold grows exponentially as a function of the total time since research in field started. Thus the rate of growth is:

- Inversely proportional with the power of the basic research function k, that is, with how difficult it is to obtain results in the field concerned
- Inversely proportional with the time ε each sample draw requires; ε can be interpreted as the time required for the production of units of basic research output that are communicable between researchers. Therefore, more frequent publication, better access to publications and improved communication technologies are all reflected in a lower value of ε.
- Proportional to the logarithm of the reduction σ of the mass of the basic research distribution by each draw.

The latter merits some further elaboration. Newton famously said he only peered further because he stood on the shoulders of giants. Our little model of learning applies this Newtonian phrase by stipulating that each new period of basic research in a field stands on the shoulders of the results of earlier periods; the parameter σ indicates how much the shoulders grow in each period. Two examples satisfying the condition of a constant reduction rate of the remaining basic knowledge distribution mass are the expected sample mean and sample maximum. In case of the sample mean we find $\sigma = 1 - (\frac{k-1}{k})^k$ and in case of the maximum $\sigma \approx 1 - \left(\frac{k-1}{k}\right)^{-k}(\frac{\varepsilon k-1}{k-1})^{-1}$. In each of these cases the threshold grows exponentially, though at different rates. In fields with rapid improvement of instrumentation due to the demand from the field itself, the value of σ is higher than in fields where no such improvement is possible.

We can now define basic research production in field x at time τ_x after the start of basic research in that field as the expected maximum sample value of the hypotheses in that field. As indicated above, we assume that the number of hypotheses that can be tested is proportional to the available amount of labor. In the appendix we show that if the labor force in a field grows exponentially, the expected sample maximum and hence the production of basic research is given by:

(11) $$\frac{dB_x}{d\tau_x} = B_0(L_x e^{\theta\tau_x})^{\frac{1}{k}}$$

Here B_0 is a constant and L_x is the basic research labor force in field x. Equation (11) represents a Harrod Domar production function with two modifications. Firstly, knowledge production is not a linear function of labor but an iso-elastic function, the elasticity being the inverse of the power of the basic knowledge distribution function. Secondly, learning is accommodated by an exponential growth of productivity in each field after research has started there. This basic knowledge production function resembles Romer's (1990) knowledge production function. The two most important differences are that (11) refers to basic knowledge only, not total technological knowledge; and that the time variable in (11) is not the absolute time of the growth model but just the time since the start of research in field x. And, of course, (11 is the result of a derivation based on explicit and fairly general assumptions about learning in basic research. Though in the derivation of (11) we assumed that the labor force in field x grows exponentially, we will also use the equation in section 5 for the case where labor force growth need not be strictly exponential.

Transitional and advanced economies

Let B_x be the supply of basic research at x, that is the cumulative production of basic knowledge pertaining to x. Then the complete condition for research driven technological progress is:

(12) $\frac{dv}{dt} > 0$ if, for $x = v,\ \delta(x,t) > \delta_c$ and $B_x \geq x$

We will assume that the actual factor productivity is the conditional expected value of the technology function. This condition makes it possible to distinguish two very different growth situations. First, suppose that the level of basic knowledge is much higher than the one corresponding to the actual technological level of a country. Then the basic research restriction in (12) is irrelevant and technological progress can be achieved by applied R and D alone. This is the situation in transitional economies such as China, India and Brazil at present, or Japan and Korea in the decades prior to 1980. Of course, in these cases, applied R and D includes the buying of licenses and other ways to import technology from abroad. Even so, however, a lot of work has to be done to master and apply technologies. Transitional economic growth is analyzed in section 4.

Targeted and free basic research

Next, in section 5, the situation of the advanced countries is analyzed, mainly North America, the European Union, Australia, New Zealand, Korea and Japan. In these countries, basic knowledge is not ahead of applied knowledge but basic research is needed to achieve technological progress.

In equation (11) the production of basic knowledge in a field depends on the labor force that is active in that field. We need to describe how this labor force in field x is determined; perhaps surprisingly, this requires some further reflection on the nature of basic research. There is little doubt that the total basic research labor force is essentially exogenous: it depends on available budgets and these, in turn, depend on politically determined public funding, on the grant positions of universities and other research institutions and on private donations. Funding of basic research is mostly separate from funding of applied R and D: the latter is largely funded by the business sector; the part that is publicly funded is mostly in areas of government services. Thus for all practical purposes the funding of basic research is separated from that of applied R and D. Consequently, we may simply assume that the basic research budget and therefore the size of the total basic research work force is exogenous; we will assume that it grows exponentially and at the same rate as the total labor force. But this leaves open how the basic research labor force is allocated to fields. At first glance it seems reasonable to assume that it is allocated proportionally to all active basic research fields. These can be defined as the fields corresponding to technology points where $v(t) < x(t) < x_c(t)$: critical experimental mass has been achieved at time t ($\delta(x,t) > \delta_c$) but the amount of basic knowledge still falls short ($B_x(t) < x(t)$). If the total basic research labor force is L_b, the allocation to fields must satisfy $L_b(t) = \int_{v(t)}^{x_c(t)} L_z dz$. A simple allocation that meets this requirement is $L_x(t) = \frac{L_b(t)}{x_c(t) - v(t)}$ and we will use this 'proportional allocation' as one of two models in our dynamic analysis.

However, this allocation is likely to be too restrictive given the nature of a basic research 'field' in our analysis. Remember that our fields are made up of a mixture of basic research disciplines as commonly understood. This means that researchers are in fact working on subjects that are not just relevant to one technology point, but to a range of them and often nobody knows where the applications will actually be. Thus, in condensed matter physics, workers are not looking at problems that are relevant to the industries where technological progress requires more basic knowledge in condensed matter physics, but on those problems where opportunities for progress in condensed matter physics appear greatest. Their solutions will not just help to remove some current barriers to technological progress in some industries, but also to help speed up innovation in other industries in the near future. In terms our x-fields corresponding to technology points, workers in all disciplines work simultaneously on all problems for a sufficient embedment in experimental and observational evidence has been achieved, irrespective of the precise applications for which this knowledge will eventually serve. Therefore, a 'joint production' model is likely to be more realistic in our description of basic research then the proportional allocation model given above. The latter assumes that the work force is completely partitioned along the lines of the technology points, whereas a joint production model assumes that the work force as whole works simultaneously on all technology points. In our dynamic analysis we will therefore use a second model which simply says $L_x = L_b$.

The difference between the proportional allocation model and the joint production model is quite interesting from a science policy point of view. In innovation policy, funding instruments are often targeted at specific economic sectors that are deemed to be of special importance or promise to a country. Academic leaders, in contrast, usually argue that basic researchers are best left free, so that they can guide themselves towards the problems where the opportunities for discovery are greatest. Our model of proportional allocation can be viewed as a representation of targeted basic research, and the joint production model as non-targeted, or free, basic research. Naturally, in the joint production model, the constant term B_0in the basic research production function must be reduced proportionally in order to obtain initial equal productivity of both models; this way, the focus of our analysis is purely on the dynamics of both models and thus on the different implications of targeted and free basic research for economic growth.

4. Growth with pure trial and error invention.

In this section we derive the growth path of neoclassical economy with the pure trial and error invention that can be associated with pre-scientific agricultural economies. We use a simple linear homogeneous Cobb-Douglas production function:

(13) $$Y = AK^{\alpha}L^{1-\alpha}\ ,\ g_Y = g_A + \alpha g_K + (1-\alpha)g_L$$

Here the growth rate of variable Z is denoted by g_Z. We assume constant rates of saving for capital accumulation and experimentation, s_K and s_D respectively, and we assume that both investments and experimentation require the same mix of goods and services as total output. Then:

(14) $$g_K = s_K\frac{Y}{K};\ g_D = s_D\frac{Y}{D}$$

Labor grows at a constant natural rate n and total factor productivity is determined by (4):

(15) $$g_L = n$$

(16) $g_A = \frac{1}{k} g_D$

Of course, if in (16) we would replace g_D by a constant, we would have Solow's standard neoclassical model with exogenous technological progress. It is now quite straightforward to derive the condition for steady growth: (13), (15) and (16) yield:

(17) $g_Y = \frac{1}{k} g_D + \alpha g_K + (1 - \alpha) n$

Denoting time derivatives by dots, (12) implies

(18) $\dot{g}_K = g_K(g_Y - g_K) \quad \dot{g}_D = g_D(g_Y - g_D)$

As $g_K$and $g_D$are strictly positive, they must converge to g_Y if the latter is constant. Substituting this into (17) yields the steady growth rate

(19) $g_Y = \frac{(1-\alpha)n}{1-\alpha-{}^{1}/_{k}}$

Thus the existence condition for steady growth is $1 - \frac{1}{k} - \alpha > 0$ or

(20) $k > \frac{1}{1-\alpha}$

Thus the power of the technology function should not be too small or, loosely formulated, innovation should not be too easy. If the power is too small, investments in trial and error generate so much technological progress and extra income that a positive feedback makes the growth rate explode (cf. figure1 below). Since such explosive growth has never been observed in pre-scientific societies, this puts a lower boundary on the power of the technology function. The lower boundary depends on α; if factor markets are efficient this coefficient corresponds with capital's factor share and (20) requires that the power of the technology exceeds the inverse of labor's share. A glance at national accounting data shows a labor share of about two-third to be common, indicating a lower boundary of 1.5 for the power. If (20) is satisfied, per capita income growth is $\frac{n}{k(1-\alpha)-1}$. This is positive if the labor force grows. At low values of k, if innovation is easy, per capita growth is highest, at higher power-values it is negligible.

To analyze the stability of the system, substitute (17) into (18) to obtain:

(21) $\dot{g}_K = g_K\{g_D \frac{1}{k} - (1 - \alpha)(g_K - n)\} \quad \dot{g}_D = g_D\{g_D(\frac{1}{k} - 1) + \alpha g_K + (1 - \alpha)n\}$

(22) $\dot{g}_K \gtreqless 0 \leftrightarrow g_K \lesseqgtr n + \frac{1}{k(1-\alpha)} g_D \qquad \dot{g}_D \gtreqless 0 \leftrightarrow g_K \gtreqless \frac{(k-1)g_D}{\alpha k} - \frac{1-\alpha}{\alpha} n$

The phase diagram corresponding to (22) is given in figure 1. The left hand side shows that if the existence condition is satisfied the system converges to the steady growth path from any initial situation; the right side shows explosive growth if the existence condition is violated.

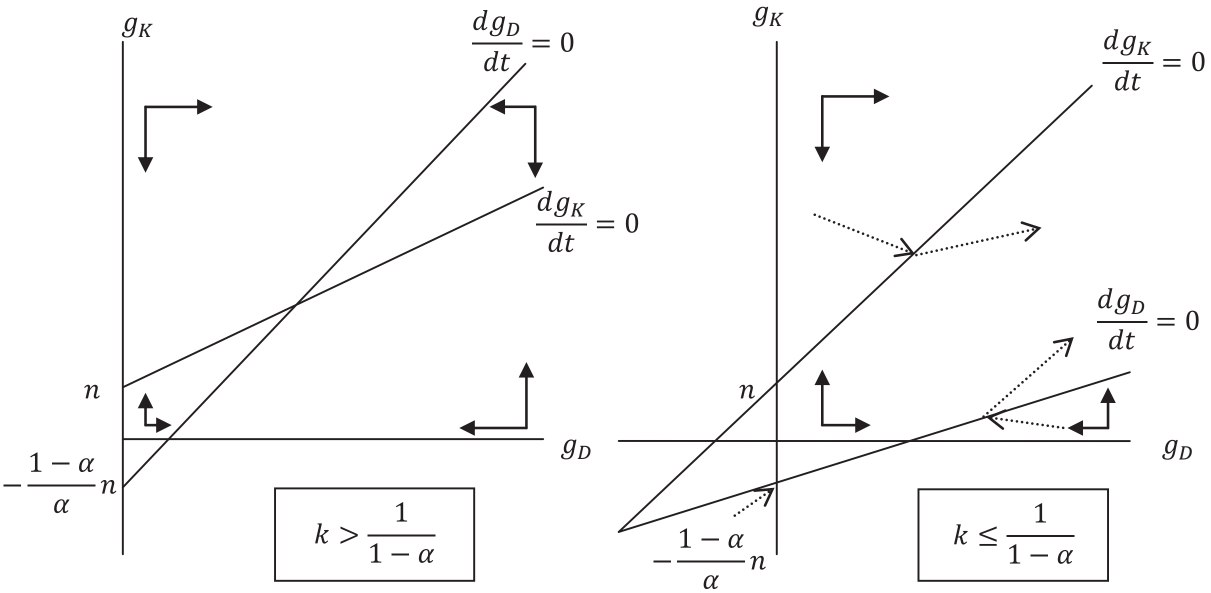

Figure 1: Phase diagrams for pure trial and error growth.

5. Transitional growth

In this section we derive the growth path of economies that invest in applied R&D but do not have to invest in basic research because the freely available level of basic knowledge is higher than that required at their level of technological development; in terms of condition (12): $B_x \geq x$. Then the threshold v in equation (6) increases to $v = x$ as soon as $\delta(x,t) \geq \delta_c$. Initially, when the cumulative experimental mass is small, the threshold is at the minimum x_{min} of the technology function. Once the experimental mass has grown sufficiently to achieve critical density at x_{min}, the threshold starts increasing and remains equal to the critical point. Until that moment $F(v) = F(x_{min}) = 0$ and (6) is simplified to $\delta(x,t) = f(x)D(t)$. Define $D_0 = \delta_c / f(x_{min})$. After criticality has been achieved at x_{min}, the threshold is equal to the critical point ($v = x_c$) and its time derivative can be obtained directly from (7):

(23) $\dot{v} = \frac{k}{k+1} \frac{1}{\delta_c} \dot{D}$

Using (5) we now obtain for the change in total factor productivity:

(24) $\dot{A} = \frac{dE(x|x \geq v)}{dt} = \frac{k}{k-1}\dot{v} = \frac{k^2}{k^2-1}\frac{1}{\delta_c}\dot{D} \quad \rightarrow \qquad \lim_{D\rightarrow\infty} g_A = g_D$

This equation replaces equation (13) of the model of section 3. The other equations, (13)-(15), remain unchanged. Since D is strictly increasing (cf. 14), (24) implies $\lim_{t\rightarrow\infty} g_A = g_D$ and (17) becomes:

(25) $\lim_{t\rightarrow\infty} g_Y = g_D + \alpha g_K + (1-\alpha)n$

Consequently, the equivalent of (21) is:

(26) $\lim_{t\rightarrow\infty} \frac{\dot{g}_K}{g_K} = g_D - (1-\alpha)\left(\lim_{t\rightarrow\infty} g_K - n\right) \qquad \lim_{t\rightarrow\infty} \frac{\dot{g}_D}{g_D} = \alpha \lim_{t\rightarrow\infty} g_K + (1-\alpha)n$

Since g_K is positive, the right hand part of (26) implies that g_K keeps on increasing as long as the natural rate is non-zero. The left hand implies

(27) $\lim_{t\rightarrow\infty} \frac{\dot{g}_K}{g_K} \gtreqless 0 \leftrightarrow \lim_{t\rightarrow\infty} g_K \lesseqgtr \frac{g_D}{1-\alpha} + n$

Thus if g_K is smaller than the extreme right side of (27) it increases, and if it larger it also increases since the extreme right hand side itself increases. It is not difficult to see that g_K and $g_D$also keep on increasing if the natural rate is zero. Consequently, with transitional growth, the growth rate of per capita income keeps on increasing and we obtain explosive economic growth. This corresponds quite well with the double digit growth usually seen in transitional economies. However, as soon as technology threshold v has increased so much that it approaches the level associated with the state of basic knowledge (formally: when $B_x \geq v$ starts to be violated), the explosive growth will end and a new regime will be entered, dictated by the advance of basic knowledge.

6. Technological progress by basic research: the steady state growth rate.

In this section we investigate how basic research as defined in section 3 generates economic growth when the level of basic knowledge is the restrictive factor in technological progress. We first give the basic equations, then analyze the existence conditions for steady growth, then the stability of the steady growth rate and finally the relation between the steady state level of income and the propensity to spend on basic research. We analyze the steady state for both for the joint production and the proportional allocation models introduced in section 3, but leave the derivation for the proportional allocation model to the appendix.

Dynamics

As in section 5, we can retain equations (13)-(15) but have to develop a new equation for the total factor productivity *A*. As in section 5, $x_c(t)$ is the *x*-value where critical mass is reached at time t. Critical mass is determined by (7) and can be written in growth rates as:

(28) $\dot{g}_x = g_x\{kg_v - (k+1)g_x + g_Y\}$

Here $g_x$is the growth rate of $x_c(t)$. The growth rate of the threshold, and therefore the rate of technological progress, is now determined by basic research. In section 3 we already introduced the assumption that the total basic research work force grows exponentially at the natural rate:

(29) $L_b = L_{B0}e^{nt}$

Because criticality is reached in field x at time *t,* this is also the time that basic research starts in field *x.* In accordance with (10) the threshold moves to *x* after the cumulative amount of basic research in this field has reached the value *x.* Let this be at time $t + \bar{\tau}_x$. Then the value of $\bar{\tau}_x$ is determined by the basic research production function (11) and the allocation of the basic research work force. For the joint production model we obtain:

(30) $x_c(t) = v(t + \bar{\tau}_x) = B_x(t + \bar{\tau}_x) = \int_t^{t+\bar{\tau}_x} B_0(L_{B0}e^{nz}e^{\theta(z-t)})^{\frac{1}{k}}dz$

Total factor productivity then is (cf. 5):

(31) $A(t + \bar{\tau}_x) = \frac{k}{k-1}v(t + \bar{\tau}_x)$

In the appendix we derive a more convenient growth rate version of (30):

(32) $(g_x + \frac{\theta}{k})x_c(t) = B_0(L_{B0}e^{nt})^{\frac{1}{k}}\{\left(1 + \frac{d\bar{\tau}_x}{dt}\right)e^{\frac{n+\theta}{k}\bar{\tau}_x} - 1\}$

If $g_{A,t+\tau_x}$ is the growth rate of A at $t + \tau_x$, (31) and the left hand side of (30) lead to:

(33) $g_{A,t+\bar{\tau}_x}\left(1 + \frac{d\bar{\tau}_x}{dt}\right) = g_x$

Together with equations (13)-(15) and (28), equations (33) and (33) give a complete description of the growth dynamics of the economy. The two new equations show that the model has a feature that is unusual in neoclassical growth theory: the role of time. In most growth models, the whole future of the system is completely determined by the current values of stocks and flows. Of course stocks are the cumulated results of past flows, but other than that there is no influence of history in the system. In our model, however, there is another role of time: there is a variable lag, like an incubation or gestation period, between the time that critical experimental mass is achieved at a technology point and the moment that this point becomes the determinant of factor productivity. This means that history matters: the future depends on both the current states of the variables and on the historical path of the critical point. The most prominent growth models that shared this feature to some extent are the vintage models of the nineteen seventies which had some promise of providing an endogenous explanation of technological progress but have fallen into disuse, probably because they are cumbersome and focus too much on the embodiment of technological progress in capital goods.

The length of the incubation period is endogenous. This is what makes steady growth possible in our model. Just as in the preceding section, the critical technology point advances ever more rapidly due to the growth of total factor productivity. In section 5, this ever more rapid advance caused an acceleration of the growth of total factor productivity and thus explosive growth. In the present section however, an increase in the gestation period between the moment that a technology point reaches criticality and the completion of basic research at that same point, counteracts the tendency to explosive growth, playing a role the similar to that of increasing product diversity in endogenous growth theories. In our case, however, the gestation period is self adjusting and there is no need for exogenous fine-tuning of some parameter.

Steady growth with joint production

The definition of steady growth is that (per capita) income grows at a constant rate, which we denote as γ. Equation (14) implies that the growth rates of capital *K* and experimental mass *D* cannot be zero if income grows and (18) that they must eventually be equal to γ. Then (13) implies $g_A = (1-\alpha)(\gamma - n)$. It is easy to see (cf. the appendix) that in the steady state the critical point equation (28) implies:

(34) $g_x = \frac{(\gamma-n)\{1+k(1-\alpha)\}+n}{k+1}$

We have now found that in the steady state the growth rates of *Y, K ,D, A* and x_c must all be constant. Of course this implies that in (33) the growth rate of *A* at time $t + \bar{\tau}_x$ is equal to that at time *t*. Therefore (31) implies:

(35) $\frac{d\bar{\tau}_x}{dt} = \frac{g_x - g_A}{g_A} = \frac{\alpha(\gamma-n)+n}{(k+1)(1-\alpha)(\gamma-n)} \equiv \rho$

Clearly, ρ must be positive; otherwise $\bar{\tau}_x$ would eventually be negative. This implies that implying that per capita income must grow at a strictly positive rate $(\gamma > n)$. In the appendix we combine (33) and (30) and find that for large enough *t* the existence of steady growth requires:

(36) $\quad kg_x - (n+\theta)\rho - n = 0$

From the three equations (34)-(36) g_x and ρ can be eliminated; this leads to quadratic equation in the growth rate of per capita income. Its solution (the negative root is ruled out because it violates the existence condition $> n$) is derived in the appendix:

(38) $\quad \gamma - n = \frac{n+\theta\alpha+\sqrt{(n+\theta\alpha)^2+4n(n+\theta)k(1-\alpha)(1+k(1-\alpha)\}}}{2k(1-\alpha)\{1+k(1-\alpha)\}}$

Though this is a complicated expression some conclusions are easy to draw. Most importantly: there is growth even if the labor force is constant ($n = 0$), provided there is a positive rate of learning in basic research ($\theta > 0$). In terms of the endogenous growth literature, this means that size-independent growth is possible. At the other end of the scale, even without learning in basic research ($\theta = 0$), there is still per capita growth as long as the natural rate is positive. Thus there are two primary sources of per capita growth: labor growth and basic research.

Some calculations show the growth rate to be a decreasing function of the power of the technology and basic knowledge function. A numerical illustration is given in figure 2. The capital elasticity of output or non-labor factor share, α, is set at the 33 percent that is commonly derived from national accounts statistics, the natural rate is one percent and the rate of learning in basic research, θ, is five percent. The figure clearly shows that a higher power law has a negative effect on per capita growth and a higher rate of learning a positive one. The influence of the learning rate is almost linear, whereas the influence of the power of the technology function itself resembles a power law. The figure suggests that actual per

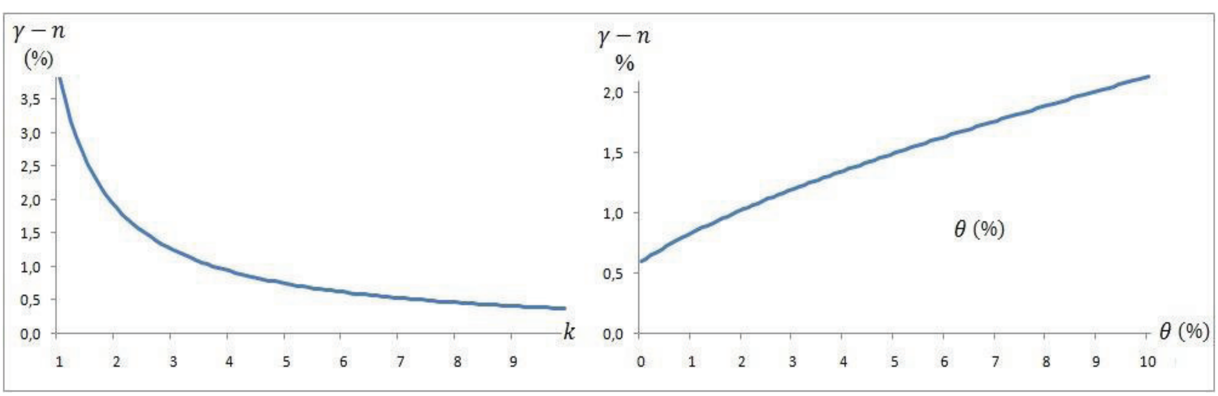

Figure 2. Impact of the technology power and basic research learning on per capita growth

capita growth is the result of two conflicting forces: the degree of difficulty of technological progress represented by the power k of the technology function and the ease of the development of basic knowledge, represented by the rate of learning θ in basic research. To illustrate this balance table 1 shows a set of combinations of k and θ that generate the same growth rate (1.5% per capita), for the same values of n and α used in figure 2.

Table 1 Combinations of k and θ that generate 1.5 percent per capita growth

k	θ (%)	k	θ (%)
1.25	0.6	4	13.0
1.5	1.3	5	20.0
1.75	2.1	7.5	43.3
2	3.0	10	75.0
2.5	5.0	12.5	115.0
3	7.3	15	163.4

$n = 1\%$ and $\alpha = 0.33$

This table shows that at a low value of the power, e.g. Lotka's two, a rate of learning of 3% is sufficient to obtain 1.5% growth; but at twice this value, 13% learning is required and above 10, the rate of learning has to be above one hundred percent. There are some situations where very high rates of learning are not implausible. One is when a field depends strongly on a technology or methodology where rapid progress occurs; ICT, with the sixty percent growth rate of computing capability of Moore's law is a case in point, but earlier examples are the inventions of logarithms, which greatly improved the productivity in all fields where computations were essential, and calculus. Even more important is the example of genomics where sequencing technology is improving at a rate that surpasses even Moore's law, in part driven by the demand for improvement from basic research itself. In general, though, lower rates of learning seem more plausible which would imply lower values of k.

Proportional allocation of the basic research labor force

In this section we consider the situation where the basic research labor force is completely specific to technology points and fully fragmented into the fields corresponding to these points. Then we need an allocation of the total basic research labor force to the technology points. In section 3 we gave a simple but consistent allocation: the labor force is allocated equally to all technology points where critical experimental mass has been achieved but more basic knowledge is needed to move the threshold. In this case the derivation is more laborious and the expression for the growth rate more complicated, but steady growth does exist as shown in the appendix.

Here we do not give the solution but only consider the difference with the case of joint production. This is shown in table 2 for various combinations of the parameter values.

Table 2 The growth rate with joint production (J) and the ratio of the growth rate with proportional allocation to that with joint production (P/J)

n	1		1		1		3		n	1		1	
α	33		33		66		33		α	33		33	
θ	5		10		5		5		k	2.5		5	
k	J	P/J	J	P/J	J	P/J	J	P/J	θ	J	P/J	J	P/J
1.25	3.0	50	4.4	53	8.3	38	5.7	38	1	0.8	60	0.4	76
1,5	2.5	56	3.6	58	6.7	44	4.8	44	2	1.0	64	0.5	79
2	1.9	63	2.7	66	4.8	53	3.6	57	4	1.3	68	0.7	82
2.5	1.5	69	2.1	71	3.8	60	2.9	63	5	1.5	69	0.7	82
4	0.9	79	1.3	80	2.2	73	1.8	73	10	2.1	71	1.0	84
5	0.7	82	1.0	84	1.7	79	1.5	78	15	2.7	72	1.3	85
10	0.4	91	0.5	92	0.8	89	0.7	88	25	3.7	73	1.7	86

In %, except *k*.

The table illustrates that the growth with proportional allocation is always lower than that with joint production. The loss is largest at low values of *k,* when innovation is easiest, and becomes negligible at high values of *k.* However, at these high values of *k* both growth rates are relatively low anyway. Similarly, a high natural rate and a high factor share of capital lead to both a high joint production growth rate and a high loss of growth in case of proportional allocation. In contrast, when the rate of learning in basic research varies, though both growth rates are highest at high rates of learning, the relative loss by proportional allocation is strongest at *low* rates of learning. Though these results seem contradictory at first glance, they are due to the same cause. Both when innovation is difficult (a high *k*) and when the rate of learning in basic is high, the impact of the growth of the basic research work force on the total growth rate is small relative to the rate of learning in basic research. In that situation, it does not matter much that the work force is allocated inefficiently.

The results with respect to proportional allocation indicate that a policy of targeting of basic research is not likely to be beneficial to economic growth but tends to lower it compared to a policy of free basic research. The loss of growth may be as much as one quarter. This conclusion might be reinforced if account is taken of the fact that targeted basic research is often targeted at sectors that are deemed to be important for society or the economy, but not at the basic knowledge fields where progress is fastest. This means that targeting may well lower the aggregate rate of learning in basic research, further increasing the loss compared to a policy of free basic research. Our model indicates that the most effective growth policies are those that enhance the joint production character of basic research: open access of publications, data, software, and large research facilities; strong research institutions with intensive interaction between people from different disciplines; rapid dissemination of early results and reliable high speed communication facilities; rapid dissemination of enabling technologies; and all other measures that allow for rapid spread of knowledge between researchers, disciplines, and institutions. Of course there is nothing new about the presumption that these measures raise productivity in basic research, in fact is a completely traditional and orthodox position. What *is* new, however, is the conclusion that these measures may lead to a higher rate of economic growth and that detracting from the joint production, open character of basic research seriously harms economic growth in advanced countries.

One proviso should be made with respect to this conclusion: strictly speaking our analysis is only valid for all advanced countries combined. In theory it is possible that individual countries might steal a march on the other advanced nations by targeting their basic research. To analyze this possibility a multi country model would be needed which analyzes both the international dissemination of research results and the transmission of growth through international trade. It should be noted, however, that such a targeted basic research policy by individual countries is inherently protectionist and must be harmful to the advanced countries combined.

7. Stability of science-based growth

In many growth models stability analysis is relatively straightforward, because the changes of the variables at a given time depend only on stocks and flows at same time. If at that time the stocks are not at their dynamic equilibrium values, flows react on the gap between current and equilibrium values, leading towards system equilibrium if the system is stable. In our case, however, the dynamics are different because of the incubation period of basic research. If the stocks are out of equilibrium, the flows adjust towards temporary equilibrium values, given the current state of technology. But this adjustment process changes the future state of technology (after the incubation period) and therefore the future temporary equilibrium values. In addition, the incubation period itself is variable. Thus there might be oscillatory or even chaotic behavior in the long term even when the system would be well behaved in the shorter term. Therefore we cannot reduce the dynamics to a two dimensional phase diagram as in figure 1. Instead employ a two stage analysis. First we analyze the adjustment of the variables towards their temporary equilibrium values or 'attractors'. Next we consider how these attractors move in the long term and approach their own equilibrium values.

The adjustment towards the temporary equilibrium is determined by the changes in the growth rates of capital and the critical technology point given the current level of technological progress; these are given by equations (15) and (28), respectively. Substitution of (13) yields:

(39) $\dot{g}_K = g_K\{g_A + (1-\alpha)(n - g_K)\}$

(40) $\dot{g}_x = g_x\{(1+k)(g_A - g_x) + \alpha g_K + (1-\alpha)n\}$

To simplify the analysis of the incubation period and the future rate of technological progress, we prove in the appendix that for large *t* the change in the incubation period approaches

(41) $\frac{d\bar{\tau}_x}{dt} = \frac{k}{n+\theta}\left(g_x - \frac{n}{k}\right)$

and that the values of g_x, g_A exceed $\frac{n}{k}$. Similarly, for large *t* the relation between the rate of technological progress at $t + \bar{\tau}_x$ and g_x approach:

(42) $g_{A,t+\bar{\tau}_x} = \frac{(n+\theta)g_x}{\theta + k g_x}$

Differentiate (42) with respect to time, substitute (40) into the result and replace g_x by the inverse of (42) to obtain the relation between the change in the rate of technological progress at $t + \bar{\tau}_x$ and the rates of technological progress and capital growth at t:

(43) $\dot{g}_{A,t+\bar{\tau}_x} \gtreqless 0 \leftrightarrow g_{A,t+\bar{\tau}_x} \lesseqgtr \omega(g_A, g_K) = \frac{(n+\theta)\{g_A(1+k)+\alpha g_K+(1-\alpha)n\}}{g_A k(1+k)+\alpha k g_K+(1-\alpha)kn+\theta(1+k)}$

In the appendix we show that ω is positive and smaller than $\frac{n+\theta}{k}$, that it is an increasing function of g_K with a decreasing slope and that it tends asymptotically to $\frac{n+\theta}{k}$ at large values of g_K. Thus a phase diagram at a given value of g_A, has the general shape given in figure 3.

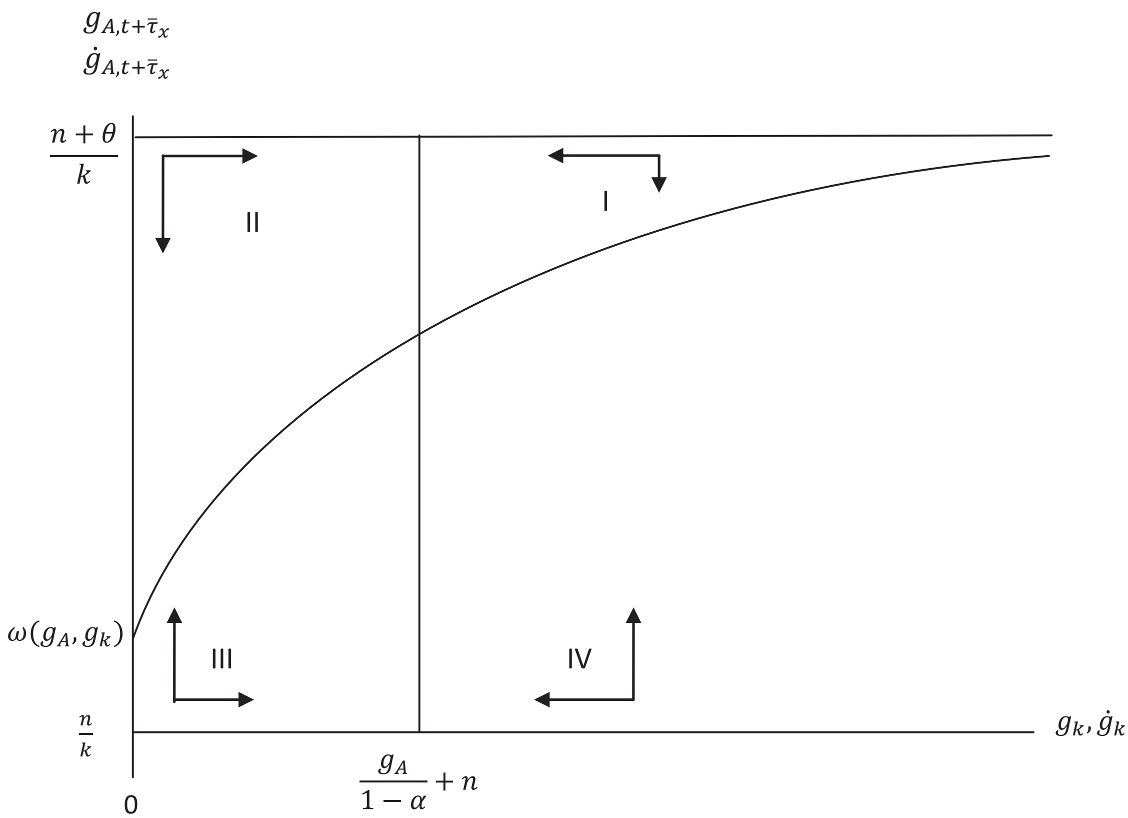

Figure 3 Phase diagram at time t for given g_A.

If the two attractors $\omega(g_A, g_k)$ and $\frac{g_A}{1-\alpha}+n$ would be static and the system would start in either quadrant I or quadrant III, the system would head directly for their intersection without crossing either of the attractors. Thus the movement of $g_{A,t+\bar{\tau}_x}$would be monotone. Similarly, if the movement of g_A is monotone between a point in time and the time when the corresponding incubation period is finished, and if at the end of that period the system is in either quadrant I (in case of a downward moving g_A) or in quadrant III (in case of an upward movement), the attractors and the variables will from that time on move monotonously in the same direction and converge to the steady state. If, however, the system starts out in either of quadrants II or IV, the attractor ω may be crossed and the direction of movement of $g_{A,t+\bar{\tau}_x}$changed. Since the future values of the attractors depend on $g_{A,t+\bar{\tau}_x}$, the attractor will not merely move in monotonously but change direction. In principle, this might imply oscillatory or chaotic behavior in the long term. To ascertain that the system nevertheless converges to the steady growth rates, we analyze the changes in the minima and maxima of the attractors.

In the appendix we show that the partial derivative of ω with respect to g_A is positive but that of $\omega - g_A$ negative; moreover, ω is larger than g_A at the minimum $g_A = \frac{n}{k}$ and smaller at the maximum $g_A = \frac{n+\theta}{k}$. In between, there is a line $\bar{\omega} = \bar{g}_A$ where both are equal. The shape of the line is similar to that of $\omega(g_A, g_k)$ in figure 3 and its intersection with the attractor $g_K = \frac{\bar{g}_A}{1-\alpha} + n$ is the steady state technological growth rate

(44) $\hat{g}_A = (1-\alpha)(\gamma - n)$,

where γ is given by (38). The movement of the attractors at some time t is determined by the values of g_A before the incubation period that ends at t. These values are constrained by the minimum $\frac{n}{k}$ and maximum $\frac{n+\theta}{k}$ of g_A. But this means, as is clear from figure 3, that if g_K is to the left of the attractor value $\frac{g_A(min)}{1-\alpha} + n$, g_K must increase until it is at least equal to this value; similarly, if g_K is to the right of the attractor value $\frac{g_A(max)}{1-\alpha} + n$, it must decrease until it is at most equal to this value. This implies that if g_A is below the ω line corresponding to $g_A(min)$, $g_{A,t+\bar{\tau}_x}$ must increase until is arbitrarily close to the intersection of that line with $g_K = \frac{g_A(min)}{1-\alpha} + n$; and when it is above the ω line corresponding to $g_A(max)$, $g_{A,t+\bar{\tau}_x}$ must decrease until it is arbitrarily close to the intersection with $g_K = \frac{g_A(max)}{1-\alpha} + n$. However, these minimum and maximum values of $g_{A,t+\bar{\tau}_x}$ are the values of g_A determining the movement of the system after the end of the incubation period. Consequently, after some point in time, the value of g_A is always above this minimum and below this maximum. This implies that there is a new, higher, minimum value for the attractor of g_K and a new maximum value. Thus we obtain a series of increasing minima and a series of decreasing maxima. This is illustrated in figure 4 for the series of minima, min, 1, 2, 3, and so on. Both series converge towards the intersection of $\bar{\omega}$ with the attractor of g_K, that is to the steady state growth rates of g_K and g_A. Therefore the steady state is globally stable.

This is an attractive property because it implies that the model and its steady state growth rate is an appropriate description of a wide range of situations. The lower left hand quadrant, III, may be thought to correspond with the gradual emergence of knowledge-led growth in the modern global economy, from the beginning of the seventeenth century in The Netherlands, diffused to England in the second half of the seventeenth century, to some other European countries and North America in the nineteenth century and to many other countries in the twentieth century. On average, the adjustment process is a gradual increase in the rates of technological progress and capital growth towards long-term equilibrium values.

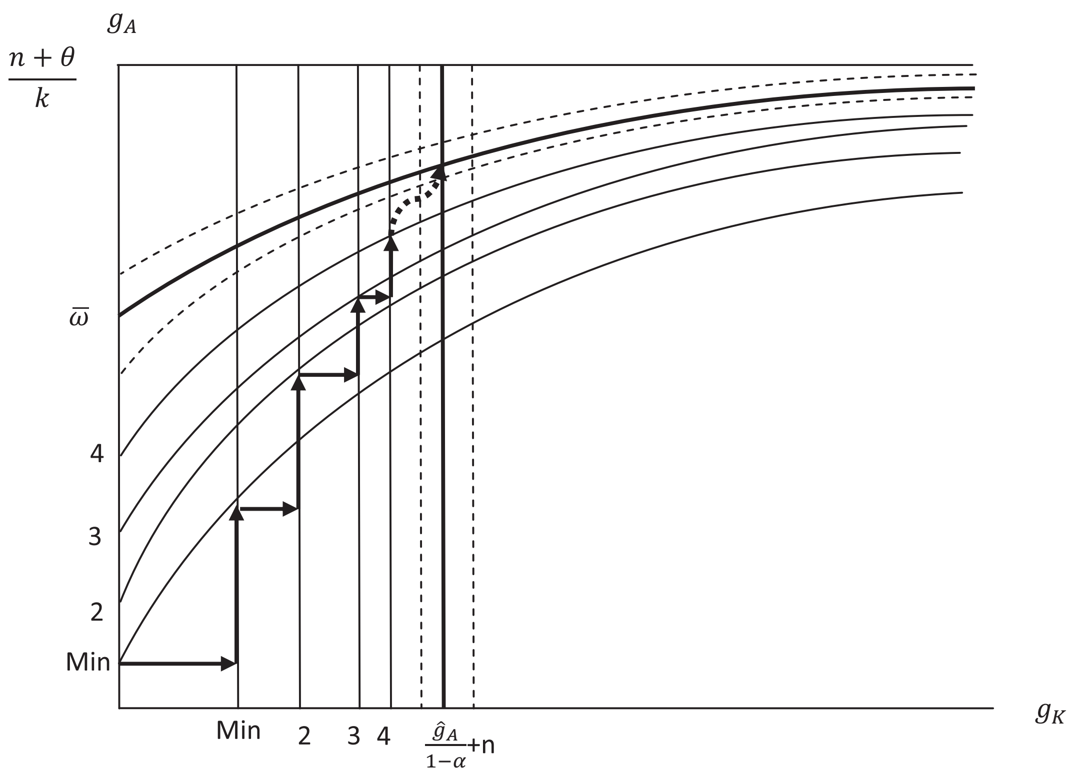

Figure 4 Development of minimum and maximum rates of technological progress g_A

The upper right-hand part quadrant I can be identified with the situation after the major global wars, both world wars in the last century and perhaps the Napoleonic wars in the early nineteenth century. During these wars basic research continued, though perhaps at a lower level of resources. Therefore the basic research learning processes were not disrupted but the application of the results was, because the amount of resources devoted to applying research results in civilian production was minimized (in the Second World War the US even ceased producing any private motor cars at all, let alone incorporate innovations). Consequently, after these periods ended there was a huge backlog of unapplied basic knowledge and the growth rates of both technology and capital are likely to have ended up in quadrant I in all nations with an advanced knowledge base. In that case our model predicts a gradual monotone convergence from above to the steady state level. This fits in quite well with the post-war booms in the nineteen twenties and the nineteen fifties and sixties; and with slower growth in the leading countries in the nineteen thirties and seventies.

The upper left hand quadrant II may be identified with the temporary divergence from long term equilibrium that is generated by the rapid growth of large transitional countries such as China, and, more recently, India and Brazil. Following our transitional growth model, the double digit growth is driven by technological knowledge as well as the rapid integration of large labor forces into the modern global economy; it closes the gap between the level of the transitional countries' total factor productivity and the productivity level consistent with the basic knowledge developed in the advanced countries. This phase is characterized by an average global rate of technological progress above the long term equilibrium; this means that the attractor of the growth rate of the global capital stock (in the sense of production equipment) rapid increases and that the actual rate of growth of capital is below it. In our model this situation causes temporary instability: there is downward influence on the future rate of technological progress from its 'too high' current global level and upward influence from the increase in the rate of capital. This may lead to some temporary oscillations until a monotone progression towards the more moderate steady state growth levels is once more established.

In the same line of reasoning, the final quadrant IV may be thought to represent the temporary imbalance after a group of large transitional countries has come close to the total productivity level of the advanced countries. At that point their rate of growth of technological progress must decline sharply, to a fraction of its value during the transition; this depresses the average rate of technological progress of the global economy and subsequently may lower the attractor of capital growth so much that it falls below its current level. This means that simultaneously the future growth rate of technology increases and that of capital falls. This too may cause some temporary oscillations before a monotone progression towards the steady state is re-established.

8. The rates of return on research spending

Since steady state growth is globally stable, long term income effects of policy measures can be obtained by considering the steady state level of income. Consequently, we are now able to address what is perhaps the most important issue in science and innovation policy: the income effects of expenditure on capital investments, applied R&D and basic research, both absolute and relative to each other. More precisely, we will analyze the effects of expenditure on these three categories on 'net income' or consumption: income after subtraction of the cost of investment and R and D. In our model, expenditure on capital outlays is determined by the rate of saving s_K, expenditure on applied R&D by s_D. We do not yet have a similar parameter for spending on basic research, since we have simply worked with a given basic research labor force. This is adequate for the analysis of growth rates, but now a more fully elaborated neoclassical model is needed. Redefine L as the total labor force, with $L = L_B + L_Y$, where L_Y is the labor input in non-research production. Let p_L be the price of labor and select units of measurement such that the price of output is 1. Standard neoclassical optimization leads to $p_L = \partial Y/\partial L_Y = (1-\alpha)Y/L_Y$. Let the basic research budget be a fixed proportion s_B of income: $p_L L_B = s_B Y$. Then

(45) $$L_B = \frac{s_B L_Y}{1-\alpha} \leftrightarrow L_B = \frac{s_B}{1-\alpha+s_B} L_0 e^{nt};\ L_Y = \frac{1-\alpha}{1-\alpha+s_B} L_0 e^{nt}$$

Following the OECD (2002) Frascati Manual on science and technology statistics we refer to s_D and s_B as (applied and basic) "research intensities"; from a strict national accounting point of view, these intensities should not be based on a definition of national income as output of goods and services Y but as Y plus the cost of basic research; this however is of no practical importance since the basic research intensity is quite small. The amounts spent on capital outlays, U_K, applied R and D, U_D, and basic research, U_B, now are $U_i = s_i Y$ $(i = K, D, B)$. Consumption is output Y less expenditure on these two categories:

(46) $$C = Y - U_K - U_D = (1 - s_K - s_D)Y$$

To obtain the consumption effects of spending on capital and research we now derive (a more detailed derivation is given in the appendix) a set of three equations for the steady state levels of the key variables Y, x_c and A. First note that in the steady state capital grows at the same rate γ as income and (14) implies $K = Y s_K/\gamma$. With this equation capital can be eliminated from the production function (13), replacing L by L_Y:

(47) $$Y = L_Y \left(\frac{s_K}{\gamma}\right)^{\frac{\alpha}{1-\alpha}} A^{\frac{1}{1-\alpha}}$$

In the steady state, the growth rate of the critical technology point is given by (34). Substitute this into the original critical point equation (7) and eliminate $\dot{D}$ using (8) and v using (31):

(48) $$(\gamma - n)\{1 + k(1-\alpha)\} + n = \frac{k s_D}{\delta_c} \left(\frac{k-1}{k}\frac{A}{x_c}\right)^k \frac{Y}{x_c}$$

In the appendix to section 7 we gave an analytic solution of the basic research equation (30), viz. (A.7.1). Asymptotically this equation approaches:

(49) $e^{\bar{\tau}_x} = [B_0(\frac{s_B}{1-\alpha+s_B}L_0)^{\frac{1}{k}}\frac{k}{n+\theta}]^{\frac{k}{n+\theta}}(x_c)^{\frac{k}{n+\theta}}\,e^{-\frac{n}{n+\theta}t}$

As the steady state growth rate of A, $\hat{g}_A$, is constant, $A(t+\bar{\tau}_x) = A(e^{\bar{\tau}_x})^{\hat{g}_A}$. Moreover, (30) and (31) imply $A(t+\bar{\tau}_x) = kx_c/(k+1)$. With (49) this yields

(50) $\frac{k}{k-1}x_c = A[B_0(\frac{s_B}{1-\alpha+s_B}L_0)^{\frac{1}{k}}\frac{k}{n+\theta}]^{\frac{k\hat{g}_A}{n+\theta}}(x_c)^{\frac{k\hat{g}_A}{n+\theta}}\,e^{-\frac{n\hat{g}_A}{n+\theta}t}$

From the three equations (47), (48) and (50) we can now successively eliminate x_c and A and obtain a solution for *Y* and *C*. Differentiating the solution for *C* with respect to the rate of saving and the two research intensities we obtain the quasi elasticities $\frac{\partial C}{C\partial s_i}$. With some arithmetic we find for the marginal consumption effects of expenditure on capital outlays, applied R& D and basic research:

(51a) $\frac{\partial C}{\partial U_K} = \frac{(1-s_K-s_D)^2\frac{\partial C}{C\partial s_K}}{1-s_D+(1-s_K-s_D)s_K\frac{\partial C}{C\partial s_K}};\ \frac{\partial C}{C\partial s_K} = \frac{\alpha\{n+\theta+k^2\hat{g}_A\}}{s_K[-\alpha(n+\theta)s_K+k\hat{g}_A\{k(1-\alpha)+1\}]} - \frac{1}{1-s_K-s_D}$

(51b) $\frac{\partial C}{\partial U_D} = \frac{(1-s_K-s_D)^2\frac{\partial C}{C\partial s_D}}{1-s_K+(1-s_K-s_D)s_D\frac{\partial C}{C\partial s_D}};\ \frac{\partial C}{C\partial s_D} = \frac{n+\theta-k\hat{g}_A}{s_D[-\alpha(n+\theta)s_D+k\hat{g}_A\{k(1-\alpha)+1\}} - \frac{1}{1-s_K-s_D}$

(51c) $\frac{\partial C}{\partial U_B} = \frac{(1-s_K-s_D)\frac{\partial C}{C\partial s_B}}{s_B\frac{\partial C}{C\partial s_B}+1};\ \frac{\partial C}{C\partial s_B} = \frac{(1-\alpha)\{(1+k)\hat{g}_A-(n+\theta+k^2\hat{g}_A)s_B\}}{(1-\alpha+s_B)s_B[-\alpha(n+\theta)+k\hat{g}_A\{k(1-\alpha)+1\}]}$

Note that the steady state rate of technological progress, $\hat{g}_A$, is given by (44) and depends only on the familiar four parameters $\alpha, k, n,$ and θ. Similarly, the 'rates of return' in (51) depend only on these four parameters and on the rate of saving and the two research intensities. The values of the rates of return are shown in table 3 for various values of the parameters. For the rate of saving we use 20%, for the applied R and D intensity 1.5% and for the basic research intensity 0.5%. These figures are in the range of actual OECD figures if we take into account that basic research mostly depends on funding by governments, private non-profit institutions and university endowments but that a significant part of these funds is spend on applied R and D. The table shows that at the assumed values of the rate of saving and research intensities the marginal returns on research expenditure are huge, and much larger than those on physical capital. The return on applied R&D is at least ten times as large as that on investment in physical capital and the return on basic research is at least three times as large as that on applied R&D. Spending one extra dollar on

Table 3 The long term marginal return (increased national consumption) on extra expenditure on physical capital, applied R&D and basic research; various values of the technology power and the rate of learning in basic research

Technology power k	Basic research learning rate θ (%)	Per capita growth $\gamma - n$(%)	Return on physical capital $\partial C/\partial U_K$	Return on applied R and D $\partial C/\partial U_D$	Return on basic research $\partial C/\partial U_B$
1.25	5	3.0	1.6	22	98
2.5	5	1.5	1.3	16	66
5	5	0.7	1.0	10	38
10	5	0.4	0.8	6	20
2,5	1	0.8	0.8	6	60
2.5	5	1.5	1.3	16	66
2,5	10	2.1	1.6	21	70
2,5	15	2.7	1.8	24	74

Returns: increase of consumption, in monetary units, due an increase in the expenditure category of one monetary unit. Natural rate one percent, capital elasticity of income one third; rate of saving 20%, applied R&D intensity 1.5%, basic research intensity 0.5%.

Table 4. Cf. table 3, various values of the rate of saving and the research intensities

Rate of saving s_K (%)	Return on physical capital $\partial C/\partial U_K$	Applied R&D intensity s_D(%)	Return on applied R and D $\partial C/\partial U_D$	Basic research intensity s_B(%)	Return on basic research $\partial C/\partial U_B$
10	3.6	1	33	0.1	334
15	2.1	1.5	22	0.33	101
20	1.3	2	16	0.5	66
25	0.8	2.5	13	0.75	44
30	0.4	3	10	1	33

Units of measurement and parameter values cf. table 3, unless otherwise indicated.

applied research generates 5-25 dollars of extra consumption, one extra dollar of basic research generates 20-100 dollars extra consumption, with a central value of 65. Naturally, the values in this table depend sensitively on the rate of saving and the two research intensities. The huge marginal returns on research are caused by the low levels of the research intensities. This is shown in table 4. Clearly, the return is extremely sensitive to the value of the intensities, much more sensitive than on the (uncertain) parameter values. For basic research this is illustrated in figure 5. The marginal consumption effect of increased spending on basic research is very high at a research intensity of a few tenths of a percent and then gradually decreases to more moderate levels. Still, even at two percent the marginal return of a further increase in spending on basic research is 20.

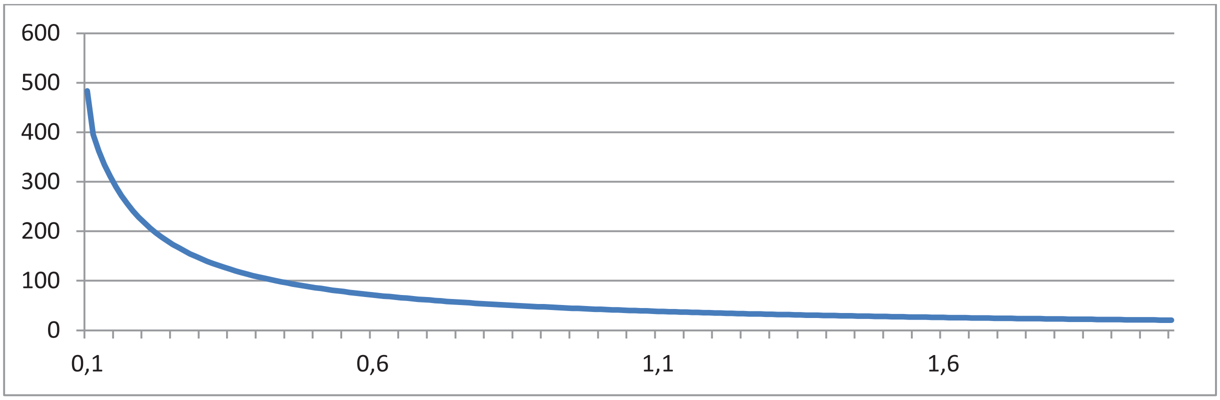

Figure 5 Return on basic research as function of basic research intensity.

The expressions in (51) imply that there are values of the rate of saving and the two research intensities that maximize steady state consumption. These 'golden rule' values are the values where the $\frac{\partial C}{C\partial s_i}$ are zero (in case of s_K and s_D both have to set at zero simultaneously):

Golden rule rate of saving: $s_K = \alpha$

Golden rule applied R& D intensity $s_D = \frac{n+\theta-k\hat{g}_A}{n+\theta+k^2\hat{g}_A}$

Golden rule basic research intensity $s_B = \frac{(1+k)\hat{g}_A}{n+\theta+k^2\hat{g}_A}$

The golden rule rate of saving is value α that is familiar from standard neoclassical growth theory. The golden values of the research intensities are quite high. For the central parameter set we used before, $\alpha = 0.33, n = 0.01, k = 2.5,$ and $\theta = 0.05$ the golden rule values of the two research intensities happen to be same: about 28.5%. These values are probably due to the lack of an inter-temporal preference rate that is an unavoidable implication of the use of a constant rate of saving and constant research intensities. It may be expected that the optimal levels of the intensities are lower if they are set to optimize an inter-temporal welfare function. However, at the values of at most a few percent presently observed in the real world, the returns on increased research spending are so huge that a realistic rate of inter-temporal preference should not significantly offset them. It is safe to conclude that rational budgetary policies in the advanced OECD countries require a substantial increase in spending on basic research and, to a lesser extent, on applied R and D.

References

Aghion, Philippe and Howitt, Peter. "A Model of Growth through Creative Destruction." *Econometrica,* March 1992, *60*(2), pp. 323-51.

--- *Endogenous growth theory.* Cambridge, MA: MIT Press, 1998.

Aghion, Philippe and Durlauf, Steven. "From Growth Theory to Policy Design." *The Commission on Growth and Development,* Working Paper No. 57*,* 2009.

Akcigit, Ufuk, Hanley, Douglas and Serrano-Velarde, Nicholas. "Back to Basics: Private and Public Investment in Basic R&D and Macroeconomic Growth." Mimeo, University of Pennsylvania, NBER, Oxford University, 2011.

Ang, James B. and Madsen, Jacob B. "Can Second-Generation Endogenous Growth Models Explain the Productivity Trends and Knowledge Production in the Asian Miracle Economies?" *Review of Economics and Statistics*, November 2011, 93 (4), pp. 1360-1373.

Berliant, Marcus and Fujita, Masahisa. "The Dynamics of Knowledge Diversity and Economic Growth." *Southern Economic Journal*, April 2011, 77(4), pp. 856-884.

Bush, Vannevar. *Science: the Endless Frontier.* Washington: Government Printing Office, 1945.

Cozzi, Guido and Galli, Silva. "Science-based R&D in Schumpeterian Growth." *Scottish Journal of political Economy*, September 2009, 56 (4), pp. 474-491.

Chu, Angus C. and Furukawa, **Yuichi** "Patentability and Knowledge Spillovers of Basic R&D." Mimeo, Durham Business School, Durham University and Shanghai University of Finance and Economics; Chukyo University, 2011.

Cozzi, Guido, Giordani, Paolo I., Zamparelli, Luca "The Refoundation of the Symmetric Equilibrium in Schumpeterian growth models." *Journal of Economic Theory*, September 2007, 136 (1), pp. 788-797.

David, H.A. *Order Statistics*. New York, London, Sydney, Toronto: John Wiley, 1970.

De Fraja, Gianni. "A theoretical Analysis of Public Funding for Research". Department of economics, University of Leicester, Working paper No. 11/31, May 2011.

Dinopoulos, Elias and Şener, Fuat. "New Directions in Schumpeterian Growth Theory." In Horst Hanusch and Andreas Pyka (eds.), *Elgar Companion to Neo-Schumpeterian Economics*, Cheltenham: Edward Elgar, 2007.

Dinopoulos, Elias and Thompson, Peter. "Schumpeterian Growth without Scale Effects." *Journal of Economic Growth*, December 1998, *3*(4), pp.313-35.

Donselaar, Piet. *Innovatie en Productiviteit: het Solow-residu Ontrafeld* (Dutch with English summary). PhD Dissertation, Rotterdam: Erasmus University, 2011.

Dosi, Giovanni, Fagiolo, Giorgio and Roventini, A. "Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles", *Journal of Economics Dynamics and Control*, 2010, *34*, pp. 1748 – 1767.

Dosi, Giovanni, Fagiolo, Giorgio, Napoletano, Mauro and Roventini, Andrea. "The Role of Technical Change, Finance and Public Policies in an Evolutionary Model of Endogenous Growth and Fluctuations." Mimeo, Paper presented at the DIME Final Conference, 6-8 April 2011, Maastricht.

Futagami, Koichi, and Iwaisako, Tatsuro. "Dynamic analysis of patent policy in an endogenous growth model." *Journal of Economic Theory,* 2007, *132*, pp. 306 – 334.

Gersbach, Hans, Schneider, Maik T. and Schneller, Olivier. "On the Design of Basic-Research Policy**."** Zürich: CER/ETH Working Paper 08/79, January 2008.

Gersbach, Hans, Sorger, Gerhard and Amon, Christian. "Hierarchical Growth: Basic and Applied Research". Zürich: CER/ETH Working Paper 09/118, August 2009.

--- "Basic Research, Openness, and Convergence." Zürich: CER/ETH Working Paper 10/139, October 2010.

Glänzel, Wolfgang. "The role of the h-index and the characteristic scores and scales in testing the tail properties of scientometric distributions." *Scientometrics* , 2010, *83, pp.* 697–709.

Griliches, Zvi. *R&D, Education, and Productivity. A retrospective*. Cambridge (MA), London: Harvard University Press, 2000.

Grossman, Gene and Helpman, Elhanan. *Innovation and growth in the global economy*. Cambridge, MA: MIT Press, 1991.

Houghton, John and Sheehan, Peter. *"*Estimating the Potential Impacts of Open Access to Research Findings.**"** *Economic Analysis & Policy*, March 2009, 39 (1), pp. 127-142.

Houthakker, Hendrik S. "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis." *The Review of Economic Studies, 1955-6, 23(1),* pp. 27-31.

Isacson, Walter. *Einstein. His Life and Universe".* New York, London, Toronto, Sydney: Simon and Schuster, 2007.

Ivancheva, Ludmila E. "The Non-Gaussian Nature of Bibliometric and Scientometric Distributions: A New Approach to Interpretation**"** *Journal Of The American Society For Information Science And Technology*, 2001, 52(13):1100–1105.

Jones, Charles I. "R&D based models of economic growth." *Journal of Political Economy*, August 1995, *103*(4), pp. 759-784.

Jones, Charles I. "Growth: with or without scale effects?" *American Economic Review*, Papers and Proceedings, May 1999, *89*(2), pp. 139-144.

Jones, Charles I. "Sources of U.S. Economic Growth in a World of Ideas." *American Economic Review*, March 2002, *92*(1), pp. 220-239.

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

Li, Wentian. "Zipf's Law Everywhere." *Glottometrics*, 2002, *5*,pp. 14-21.

Madsen, Jakob B. "The anatomy of growth in the OECD since 1870." *Journal of Monetary Economics,* 2010, *57*, pp. 753–767.

Madsen, Jakob B., Saxena, Shishir, and Ang, James B. "The Indian growth miracle and endogenous growth." *Journal of Development Economics*, 2010, *93,* pp.37–48.

OECD. *The measurement of scientific and technological activities*, *Frascati Manual.* Paris: OECD, 2002.

Lotka, Alfred J. "The frequency distribution of scientific productivity." *Journal of the Washington Academy of Sciences,* June 1926, 16(12), pp. 317-323.

Peretto, Pietro. "Technological Change and Population Growth." *Journal of Economic Growth*, December 1998, *3*(4), pp. 283-311.

Poiter, William G. "Lotka's Law Revisited."*Library trends*, Summer 1981, 30(1), pp. 21-39.

Romer, Paul M. "Endogenous technological change." *Journal of Political Economy*, October 1990, *98*(5), pp. S71-S102.

Romer, Paul M. "The origins of Endogenous Growth." *Journal of Economic Perspectives*, Winter 1994, *8*(1), pp. 3-22.

Romer, David. *Advanced Macro-economics.* Boston: McGraw-Hill, 2001.

Schies, D. and Wehrli, R. "Long-Term Growth Driven by a Sequence of General purpose Technologies." Zürich: CER/ETH, Working Paper 11/148, July 2011.

Segerstrom, Paul. "Endogenous Growth Without Scale Effects." *American Economic Review*, December 1998, *106*(1), pp. 41-63

Solow, Robert M. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics,* February 1956, *70* (1), pp. 65-94.

--- "Technical Change and the Aggregate Production Function." *The Review of Economics and Statistics*, August 1957, *39* (3), pp 312-320.

Steedman, Ian. "On 'Measuring' knowledge in New (Endogenous) Growth Theory." Paper presented at the Growth theory Conference, Pisa, Italy, October 5-7 2001. Mimeo, Department of Economics, Manchester Metropolitan University.

Van Raan, Anthony. "On growth, ageing, and fractal differentiation of science." *Scientometrics*, 2000, *47*, 347-362.

Young, Alwyn. "Growth without scale effects." *Journal of Political Economy*, February 1998, *106*(1), pp. 41-63.

Mathematical appendix

For every section of the main text this appendix gives the mathematical derivations, if needed.

Section 2, equation 4

The expected value $E(Y_r)$ of the *r*-th order statistic of a sample of size *m* from a distribution with cdf *F(x)* and pdf *f(x)* is (David, 1970, p. 25):

(A.2.1) $$E(Y_r) = m\binom{m-1}{r-1} \int_{-\infty}^{\infty} x[F(x)]^{r-1}[1-F(x)]^{m-r}f(x)dx$$

The sample maximum is the *m*-th order statistic. Therefore its expected value, $E(\hat{X}_m)$, is:

(A.2.2) $$E(\hat{X}_m) = E(Y_n)n\int_{-\infty}^{\infty} x[F(x)]^{m-1}f(x)dx$$

For the Pareto distribution with minimum 1 this boils down to

(A.2.3) $$E(\hat{X}_m) = E(Y_n) = mk\int_1^{\infty}(1-x^{-k})^{m-1}x^{-k}dx$$

Define $z = x^{-k}$, so that $zdx = -\frac{1}{k}z^{-\frac{1}{k}-1}dz$ and

(A.2.4) $$E(\hat{X}_m) = -m\int_1^0(1-z)^{m-1}z^{-\frac{1}{k}}\,dy = m\int_0^1(1-z)^{m-1}z^{-\frac{1}{k}}\,dz$$

The integral is the beta function $B\left(\frac{k-1}{k}, m\right)$; consequently:

(A.2.4) $$E(\hat{X}_m) = m\frac{\Gamma\left(\frac{k-1}{k}\right)\Gamma(m)}{\Gamma(\frac{k-1}{k}+m)} = m!\frac{\Gamma\left(\frac{k-1}{k}\right)}{\Gamma(\frac{k-1}{k}+m)} = m!\frac{1}{\prod_{i=0}^{m-1}(m-i-\frac{1}{k})} = \frac{m!k^m}{\prod_{j=1}^{m}(kj-1)}$$

If the minimum is $x_{min,}$ the expectation is given by (4a). To achieve tractability in the framework of a growth model, we switch to a continuous sample size. Note that (A.2.1) can be written as

(A.2.5) $$\ln E(\hat{X}_m) = \ln x_{min} + m\ln k + \sum_{j=1}^{m}\ln j - \sum_{j=1}^{m}\ln(kj-1)$$

Let $D \geq 1$ be the continuous sample mass. Then the equivalent of (4.a) is

(A.2.6) $$\ln E(\hat{X}) = \ln x_{min} + D\ln k + \int_1^D \ln z\,dz - \int_1^D \ln(kz-1)\,dz$$

Note

(A.2.7) $$\int \ln(kz-1) = \left(z-\frac{1}{k}\right)\ln(kz-1) - z$$

Hence:

(A.2.8) $$\ln E(\hat{X}) = \ln x_{min} + D\ln k + \{z\ln z - z\}|_1^D - \left\{\left(z-\frac{1}{k}\right)\ln(kz-1) - z\right\}|_1^D =$$

$$= \ln x_{min}(k-1) + D\ln\frac{Dk}{Dk-1} + \frac{1}{k}\ln\frac{(kD-1)}{k-1}$$

(A.2.9) $$E(\hat{X}) = x_{min}(k-1)(\frac{Dk}{Dk-1})^D(\frac{kD-1}{k-1})^{1/k}$$

(A.2.10) $\frac{d\ln E(\hat{X})}{dD} = \ln\frac{Dk}{Dk-1}$; $\frac{d\ln E(\hat{X})}{d\ln D} = D\ln\frac{Dk}{Dk-1}$

Consider $\lim_{D\to\infty} D\ln\frac{Dk}{Dk-1}$. Let $\Delta = \frac{1}{D}$, $h(\Delta) = \ln\frac{k}{k-\Delta}$ and let $h^{(i)}$ be the $i-$th derivative of h. Then a Taylor expansion of h around $\Delta = 0$ with Lagrange's remainder term yields $h(\Delta) = \frac{\Delta}{k} + \frac{\Delta^2}{2(k-\xi)^2}$ $(0 < \xi < \Delta;\ \lim_{\Delta\to 0}\xi = 0)$. Thus

(A.2.11) $\lim_{D\to\infty}\{D\ln\frac{Dk}{Dk-1}\} = \frac{1}{k}$

(A.2.12) $\lim_{D\to\infty} E(\hat{X}) = x_{min}\{(k-1)^{k-1}ekD\}^{\frac{1}{k}}$; $\frac{d\ln E(\hat{X})}{d\ln D} = \frac{1}{k}$

Section2, equation 7

A necessary condition for the threshold to increase beyond any value of x is that critical density is achieved there. Before the moment $t_0 > 0$ when the critical density is reached at x_{min} (that is $\delta(x_{min}, t_0) = \delta_c$) the threshold is at x_{min} and $F(v(t)) = F(x_{min}) = 0$. Consequently we obtain from the right hand part of (6) for $t \le t_0$

(A.2.13) $D(t_0) = \frac{\delta_c}{f(x_{min})}$

Provided D increases with time, the left hand part of (6) implies that the density increases monotonously at all values of x; the right hand part implies that at any given time, the density decreases with x, since $f(x)$ is a decreasing function. Consequently, at any moment there is only one value of x, denoted by $x_c(t)$, where $\delta(x_c(t)) = \delta_c$. Suppose the threshold remains at x_{min} until $t_1 \ge t_0$ and let $D(0)$ be negligible. Then the right-hand part of (6) can be written as:

(A.2.14) $\delta(x,t) = f(x)D(t_1) + \int_{t_1}^{t}\frac{f(x)}{1-F(v(z))}\frac{\partial D(z)}{\partial z}dz$

Insert $x = x_c(t)$ and differentiate with respect to time:

(A.2.15) $\delta_c = \delta(x_c(t)) = f(x_c(t))D(t_1) + f(x_c(t))\int_{t_1}^{t}\frac{1}{1-F(v(z))}\frac{\partial D(z)}{\partial z}dz$

(A.2.16) $0 = D(t_1)\frac{df(x_c(t))}{dt} + \frac{df(x_c(t))}{dt}\int_{t_1}^{t}\frac{1}{1-F(v(z))}\frac{\partial D(z)}{\partial z}dz + \frac{f(x_c(t))}{1-F(v(t))}\dot{D}$

Using (A.2.14) to eliminate the integral from (A.2.16) we obtain:

(A.2.17) $\frac{df(x_c(t))}{dt} = \frac{df(x_c(t))}{d(x_c(t))}\frac{dx_c(t)}{dt} = -\frac{1}{\delta_c}\frac{f^2(x_c(t))}{1-F(v(t))}\dot{D}$

From this (7) is obtained by using the Pareto function (2).

Section 3, equation 11

Let $E(\widehat{H},\tau)$ be the expected maximum of a sample of hypotheses h in the sample mass in field x. We omit the field subscript. The conditional distribution at time τ of any sample above some value of h is independent of the value of the threshold at the time that the original sample was drawn. Therefore we can apply the equivalent of (A.2.9):

(A3.1) $E(\widehat{H},\tau) \approx v(\tau)\{(k-1)^{k-1}keD(\tau)\}^{\frac{1}{k}}$

Here $D(\tau)$ and $v(\tau)$ are, respectively, the cumulative sample size and threshold at time τ. The change in the sample mass is

(A.3.2) $$\frac{dD(\tau)}{d\tau} = L - \frac{\partial D(\tau)}{\partial v(\tau)}\frac{dv(\tau)}{d\tau} = L - g\big(v(\tau)|h \geq v(\tau)\big)D(\tau)\frac{dv(\tau)}{d\tau} = L - k\theta D(\tau)$$

Now consider the case where the labor force grows exponentially: $L = L_0 e^{n\tau}$. Then the solution of the differential equation (A.3.2) is:

(A.3.3) $$D(\tau) = \frac{L_0}{n+k\theta}e^{n\tau} + e^{-k\theta\tau}$$

Substituting this into (A.3.2) yields, for large enough τ:

(A.3.4) $$E(\widehat{H},\tau) = v(\tau)\{(k-1)^{k-1}\frac{k}{n+k\theta}e\}^{\frac{1}{k}}L^{\frac{1}{k}}$$

Thus we can write

(A.3.5) $$\frac{dE(\widehat{H},\tau)}{d\tau} = B_0(Le^{\theta\tau})^{\frac{1}{k}},\ B_0 = \frac{\theta+n}{k}h_{min}\{(k-1)^{k-1}\frac{k}{n+k\theta}e\}^{\frac{1}{k}}$$

Section 6, equation 32

First note that we can write (30) as

(A.6.1) $$x_c(t) = B_0(L_{B0}e^{-\theta t})^{\frac{1}{k}}\int_t^{t+\bar{\tau}_x} e^{\frac{n+\theta}{k}z}dz$$

This equation can be solved analytically, but we use a different derivation that is also applicable in the case of proportional allocation where an analytical solution is not available. Differentiate (30) with respect to *t* using the standard formula for differentiating an integral (Courant, 1936, p. 220) and next substitute $x_c(t)$ for the right hand side of (30):

(A.6.2) $$\frac{dx_c(t)}{dt} = \frac{-\theta}{k}B_0(L_{B0}e^{-\theta t})^{\frac{1}{k}}\int_t^{t+\bar{\tau}_x} e^{\frac{n-\theta}{k}z}dz + B_0(L_{B0}e^{-\theta t})^{\frac{1}{k}}\Big\{-e^{\frac{n+\theta}{k}t} + \Big(1 + \frac{d\bar{\tau}_x}{dt}\Big)e^{\frac{n+\theta}{k}(t+\bar{\tau}_x)}\Big\}$$

(A.6.3) $$\frac{dx_c(t)}{dt} = \frac{-\theta}{k}x_c(t) + B_0(L_{B0}e^{nt})^{\frac{1}{k}}\{\Big(1+\frac{d\bar{\tau}_x}{dt}\Big)e^{\frac{n+\theta}{k}\bar{\tau}_x} - 1\}$$

Substitute $\frac{dx_c(t)}{dt} = x_c(t)g_x$ to obtain (32).

Section 6 Equation 34

Since $g_A = g_v$, the critical point equation (28) yields

(A.6.4) $\dot{g}_x = g_x\{k(1-\alpha)(\gamma-n)-(k+1)g_x+\gamma\}$ and in the steady state

(A.6.5) $\dot{g}_x \gtreqless 0 \leftrightarrow g_x \lesseqgtr \frac{(\gamma-n)\{1+k(1-\alpha)\}+n}{k+1}$

This implies that in the steady state $\lim_{t\to\infty} g_x$is given by (34).Note that this is independent of the assumptions about the allocation of the basic research labor force.

Section 6, equation 36 and 37

Differential equation (33) can be solved to obtain

(A.6.6) $\bar{\tau}_x(t) = \bar{\tau}_x(t_0) + \rho(t-t_0)$ or $\bar{\tau}_x = \bar{\tau}_0 + \rho t$

Here t_0 is a value of *t* where the steady state exists and $\bar{\tau}_0 \equiv \bar{\tau}_x(t_0) - \rho t_0$. Since g_x is constant we can write $x_c(t) = x_{c0}\, e^{g_x t}$ in the steady state. Therefore, using (35) and (A.6.6) we obtain from (32) for the steady state:

(A.6.7) $\left(g_x + \frac{\theta}{k}\right) x_{c0}\, e^{g_x t} = B_0(L_{B0}e^{nt})^{\frac{1}{k}}\left\{(1+\rho)e^{\frac{n+\theta}{k}(\bar{\tau}_0+\rho t)} - 1\right\}$

Since ρ is positive, the term minus one between curly brackets vanishes for large t and (A.6.7) implies:

(A.6.8) $e^{\{\frac{n+\theta}{k}\rho+\frac{n}{k}-g_x\}t} = \left(g_x + \frac{\theta}{k}\right)\frac{x_{c0}}{B_0(1+\rho)}L_{B0}^{-\frac{1}{k}}e^{-\frac{n+\theta}{k}\bar{\tau}_0}$

Since the right hand side is constant, this equation can only be satisfied if the exponent in the left hand side is zero, implying (36). Substitution of (34) and (35) in (36) now yields a quadratic equation in $\gamma - n$:

(A.6.9) $(\gamma-n)^2 k(1-\alpha)\{1+k(1-\alpha)\} - (\gamma-n)(n+\theta\alpha) - n(n+\theta) = 0$

The solution of this equation is (37). The negative root can be omitted: it leads to a negative value of $\gamma - n$ and is therefore, as shown above, inconsistent with steady growth. The derivatives of this growth rate are:

(A.6.10) $\frac{d(\gamma-n)}{dk} = -\frac{(n+\alpha\theta)(\gamma+\alpha\theta)\{1+2k(1-\alpha)\}}{k(1-\alpha)\{1+k(1-\alpha)\}}$

Section 6 Proportional allocation.

In case of proportional allocation, (30) is replaced by

(30b) $x_c(t) = v(t+\bar{\tau}_x) = B_x(t+\bar{\tau}_x) = \int_t^{t+\bar{\tau}_x} B_0\left(\frac{L_{B0}e^{nz}}{x_c(z)-v(z)}e^{\theta(z-t)}\right)^{\frac{1}{k}} dz$

Instead of (A.6.1) we obtain:

(A.6.1a) $x_c(t) = B_0(L_{B0}e^{-\theta t})^{\frac{1}{k}}\int_t^{t+\bar{\tau}_x}\left(\frac{e^{(n+\theta)z}}{x_c(z)-v(z)}\right)^{\frac{1}{k}}dz$

Differentiation with respect to t and substitution into the result of $x_c(t)$ for the expression equal to the right hand side of (A.6.1a) yields

(32b) $$\left(g_x + \frac{\theta}{k}\right) x_c(t) = B_0 (L_{B0} e^{nt})^{\frac{1}{k}} \{\left(1 + \frac{d\bar{\tau}_x}{dt}\right) (\frac{e^{(n+\theta)\bar{\tau}_x}}{x_c(t+\bar{\tau}_x) - v(t+\bar{\tau}_x)})^{\frac{1}{k}} - (\frac{1}{x_c(t) - v(t)})^{\frac{1}{k}}\}$$

Equations (30), (33)-(35) and (A.6.4) are unchanged from joint production. In the steady state $x_c(t) = x_{c0}\, e^{g_x t}$. As $\bar{\tau}_x = \bar{\tau}_0 + \rho t$, we have $x_c(t + \tau_x) = x_{c0} e^{g_x(\bar{\tau}_0 + \rho t + t)}$. Moreover, $v(t + \bar{\tau}_x) = x_c(t) = x_{c0}\, e^{g_x t}$. To obtain $v(t)$, define t_1 so that $t = t_1 + \bar{\tau}_x(t_1) = \bar{\tau}_0 + (\rho + 1)t_1$, or $t_1 = \frac{t - \bar{\tau}_0}{1+\rho}$. Thus $v(t) = x_c(t_1) = x_{c0}\, e^{g_x \frac{t-\bar{\tau}_0}{1+\rho}}$. The equivalent of (A.6.7) now is:

(A.6.7a) $$\left(g_x + \frac{\theta}{k}\right) (x_{c0} e^{g_x t})^{\frac{k+1}{k}} = B_0 (L_{B0} e^{nt})^{\frac{1}{k}} \{(1+\rho)(\frac{e^{(n+\theta-g_x)(\bar{\tau}_0+\rho t)}}{1 - e^{-g_x(\bar{\tau}_0+\rho t)}})^{\frac{1}{k}} - (\frac{1}{1 - e^{-g_x \frac{t\rho+\bar{\tau}_0}{1+\rho}}})^{\frac{1}{k}}\}$$

Since ρ and g_x are positive, the denominators of the terms of the expression between curly brackets approach 1 for large t and thus the expression itself approaches $(1 + \rho)e^{\{\frac{n+\theta}{k} - g_x\}(\bar{\tau}_0 + \rho t)}$. Therefore steady growth requires $g_x \frac{k+1}{k} = \frac{n}{k} + \frac{\rho}{k}\{n + \theta - g_x\}$ and the equivalent of (36) is

(36a) $$g_x(\rho + k + 1) - n(\rho + 1) - \theta\rho = 0$$

Substitute (34) and (35) into (36a):

(A.6.9a) $$(\gamma - n)^2\{1 + k(1-\alpha)\}\{1 + (1-\alpha)k(k+2)\} + (\gamma - n)[nk(1-2\alpha) + n - \theta\alpha(k+1)] - n\{nk + \theta(k+1)\} = 0$$

(38a)

$$\gamma - n = \frac{-nk(1-2\alpha) - n + \theta(k+1)\alpha + \sqrt{(nk(1-2\alpha) - n + \theta(k+1)\alpha)^2 + 4n\{nk+\theta(k+1)\}\{1+k(1-\alpha)\}\{1+(1-\alpha)k(k+2)\}}}{2\{1+k(1-\alpha)\}\{1+(1-\alpha)k(k+2)\}}$$

Section 7, equations 41 and 42

In the case of steady growth analysis, we differentiated equation (A. 6.1); for stability analysis it is more convenient to solve it directly and then differentiate the solution:

(A.7.1) $$x_c(t) = B_0 L_{B0}^{\frac{1}{k}} \frac{k}{n+\theta} e^{\frac{n}{k}t}(e^{\frac{n+\theta}{k}\bar{\tau}_x} - 1)$$

(A.7.2) $$g_x = \frac{n}{k} + \frac{n+\theta}{k} \frac{e^{\frac{n+\theta}{k}\bar{\tau}_x}}{e^{\frac{n+\theta}{k}\bar{\tau}_x} - 1} \frac{d\bar{\tau}_x}{dt}$$

Thus

(A.7.3) $$\frac{d\bar{\tau}_x}{dt} = \frac{k}{n+\theta}\left(g_x - \frac{n}{k}\right)(1-\mu),\ \mu = e^{-\frac{n+\theta}{k}\bar{\tau}_x}\ \ (0 < \mu < 1)$$

Clearly, if $g_x > \frac{n}{k}$ for all t above some value, $\bar{\tau}_x$keeps on increasing and $\lim_{t\to\infty} \mu = 0$. In that case we can write:

(A.7.4) $$\frac{d\bar{\tau}_x}{dt} = \frac{k}{n+\theta}\left(g_x - \frac{n}{k}\right) + \Delta;\ \lim_{t\to\infty} \Delta = 0$$

Using this equation and (33) the asymptotic relation (42) between $g_{A,t+\bar{\tau}_x}$and $g_x$is obtained. It remains to prove that $g_x > \frac{n}{k}$ is indeed satisfied above some value of *t*. First note that (39) implies

(A.7.5) $$\dot{g}_K > 0 \text{ if } g_K < \frac{g_A}{1-\alpha} + n \text{ and } g_K \geq n - \Delta, \lim_{t\to\infty} \Delta \leq 0$$

Note that we use the same Δ as in (A.7.4). This is valid here as well as in the next few equations since Δ is just a remainder term that is arbitrarily small if *t* is chosen sufficiently large. Now (40) implies

(A.7.6) $$\dot{g}_x \geq g_x\{-(1+k)g_x + n - \Delta\} \text{ and } g_x \geq \frac{n}{1+k} - \Delta$$

Thus after some time, g_K and g_x have minimum values close to n and $\frac{n}{k+1}$, respectively. Next, substitute (A.7.3) in (33):

(A.7.7) $$g_{A,t+\bar{\tau}_x} = \frac{g_x(n+\theta)}{\theta + kg_x + \mu(n-kg_x)} \qquad g_x = \frac{(\theta+\mu n)g_{A,t+\bar{\tau}_x}}{n+\theta-k(1-\mu)g_{A,t+\bar{\tau}_x}}$$

Note that the denominator of $g_x$must be non negative, since $g_x$is non-negative. This requires

(A.7.8) $$g_{A,t+\bar{\tau}_x} \leq \frac{n+\theta}{k(1-\mu)}$$

Moreover, substituting the extreme right hand side of (A.7.7) into $g_x \geq \frac{n}{1+k} - \Delta$, we obtain after some reshuffling $g_{A,t+\bar{\tau}_x} \geq [n - \Delta_3(1+k)]/[k + \frac{\theta+\mu n}{\theta+n} - \Delta_3\frac{(1+k)k(1-\mu)}{\theta+n}]$. Since $\frac{\theta+\mu n}{\theta+n} \leq 1$ this implies $g_{A,t+\bar{\tau}_x} \geq \frac{n}{1+k} - \Delta$. Consequently, from some time on, the value of $g_A$satisfies (A.7.9). Applying reasoning similar to that in (A.7.5) we now obtain after some point in time:

(A.7.10) $$g_K \geq \frac{n}{(1-\alpha)(1+k)} + n - \Delta$$

Now (40) implies $\dot{g}_x \geq g_x\left\{n\{2 + \frac{\alpha n}{(1-\alpha)(1+k)}\} - (1+k)g_x - \Delta\right\}$. Because $k \geq 1$we have $\frac{n}{1+k}\left\{2 + \frac{\alpha n}{(1-\alpha)(1+k)}\right\} \geq \frac{n}{k} + \frac{\alpha n}{(1-\alpha)(1+k)^2}$. Therefore (A.7.11) implies $g_x \geq \frac{n}{k} + \frac{\alpha n}{(1-\alpha)(1+k)^2} - \Delta$. As Δ is arbitrarily close to zero if t is large enough, this implies that after some point in time g_x is indeed strictly larger than $\frac{n}{k}$.

Section 7 Equation (43) and figure 3

Differentiate (42) with respect to time:

(A.7.11) $\dot{g}_{A,t+\bar{\tau}_x} = \frac{\theta \dot{g}_x}{[\theta+kg_x]^2} = \frac{\theta g_x\{(1+k)(g_A-g_x)+\alpha g_K+(1-\alpha)n\}}{[\theta+kg_x]^2}$

Use (42) to write g_x as a function of $g_{A,t+\bar{\tau}_x}$:

(A.7.12) $g_x = \frac{\theta g_{A,t+\bar{\tau}_x}}{n+\theta-kg_{A,t+\bar{\tau}_x}}$ ($g_{A,t+\bar{\tau}_x} \leq \frac{n+\theta}{k}$ is implied by $g_x \geq 0$)

Substitute this into (A.7.11):

(A.7.13) $\frac{\dot{g}_{A,t+\bar{\tau}_x}}{g_{A,t+\bar{\tau}_x}} = \frac{-g_{A,t+\bar{\tau}_x}[(1+k)\theta+k(1+k)g_A+\alpha kg_K+(1-\alpha)kn]+\{n+\theta\}\{(1+k)g_A+\alpha g_K+(1-\alpha)n\}}{[n+\theta]^2}$

This implies

(A.7.14) $\dot{g}_{A,t+\bar{\tau}_x} \gtreqless 0 \leftrightarrow g_{A,t+\bar{\tau}_x} \lesseqgtr \frac{(n+\theta)\{g_A(1+k)+\alpha g_K+(1-\alpha)n\}}{g_Ak(1+k)+\alpha kg_K+(1-\alpha)kn+\theta(1+k)} = \omega(g_A, g_K)$

In order to sketch the phase diagram in figure 3 we need to derive the properties of ω. First note that ω is positive for all values of g_A, g_K. Moreover, it is smaller than $\frac{n+\theta}{k}$:

(A.7.15) $\omega - \frac{n+\theta}{k} = \frac{n+\theta}{k}\frac{-\theta(1+k)}{g_Ak(1+k)+\alpha kg_K+(1-\alpha)kn+\theta(1+k)} < 0$

This equation immediately implies $\lim_{g_K\to\infty}\omega = \frac{n+\theta}{k}$. Next consider the partial derivative to g_K

(A.7.16) $\frac{\partial\omega}{\partial g_K} = \frac{\theta(1+k)(n+\theta)\alpha}{[k\{g_A(1+k)+\alpha g_K+(1-\alpha)n\}+\theta(1+k)]^2}$

This is positive and decreases as g_K increases. Furthermore

(A.7.17) $\omega(g_K=0) = \frac{(n+\theta)\{g_A(1+k)+(1-\alpha)n\}}{k\{g_A(1+k)+(1-\alpha)n\}+\theta(1+k)} > 0$

(A.7.18) $\omega\left(g_K = \frac{g_A}{1-\alpha}+n\right) = \frac{(n+\theta)\{g_A(1+k-k\alpha)+n(1-\alpha)\}}{g_Ak(1+k-k\alpha)+kn(1-\alpha)+\theta(1+k)(1-\alpha)}$

This allows the sketch of the short term phase diagram, for given g_A, in Figure 3.

Section 7, Figure 4

We first give a number of properties of ω as a function of g_A and then derive figure 4.

(A.7.19) $\frac{\partial\omega}{\partial g_A} = \frac{(n+\theta)\theta(1+k)^2}{[g_Ak(1+k)+\alpha kg_K+(1-\alpha)kn+\theta(1+k)]^2} > 0$

Thus ω is an increasing function of g_A. The difference between ω and g_A is

(A.7.20) $\omega - g_A = \frac{-g_A^2k(1+k)+g_A[n-\alpha kg_K+\alpha kn]+(n+\theta)\{\alpha g_K+(1-\alpha)n\}}{g_Ak(1+k)+\alpha kg_K+(1-\alpha)kn+\theta(1+k)}$

The partial derivative of this function with respect to g_A is

(A.7.21) $\frac{\partial(\omega-g_A)}{\partial g_A} = \frac{(n+\theta)\theta(1+k)^2-[g_Ak(1+k)+\alpha kg_K+(1-\alpha)kn+\theta(1+k)]^2}{[g_Ak(1+k)+\alpha kg_K+(1-\alpha)kn+\theta(1+k)]^2}$

The sign of this derivative is the sign of

(A.7.22) $$(1+k)\sqrt{(n+\theta)\theta} - g_A k(1+k) - \alpha k g_K - (1-\alpha)kn - \theta(1+k)$$

At the minimum of g_A, $\frac{n}{k}$, this expression has the value

(A.7.23) $$(1+k)\sqrt{(n+\theta)}\{\sqrt{\theta} - \sqrt{(n+\theta)}\} - nk(1-\alpha) - \alpha k g_K < 0$$

Thus $\omega - g_A$is a strictly decreasing function of g_A. The value of $\omega - g_A$ is zero at

(A.7.24) $$\bar{\omega} = \bar{g}_A = \frac{[n-\alpha k g_K+\alpha kn]+\sqrt{[n-\alpha k g_K+\alpha kn]^2+4k(1+k)(n+\theta)\{\alpha g_K+(1-\alpha)n\}}}{2k(1+k)}$$

With respect to the sign of $\omega - g_A$ we now have

(A.7.25) $$\omega(g_A, g_K) - g_A \gtreqless 0 \leftrightarrow g_A \lesseqgtr \bar{g}_A(g_K)$$

The intersection of the numerator of $\omega - g_A$ with $g_K = \frac{g_A}{1-\alpha} + n$ is

(A.7.26) $$\omega\left(g_K = \frac{g_A}{1-\alpha} + n\right) - g_A = \frac{-g_A^2 k[1+k-\alpha k]+g_A[n+\theta\alpha]+(n+\theta)n(1-\alpha)}{g_A[1+k-\alpha k]+(1-\alpha)\{kn+\theta(1+k)\}}$$

This is zero if

(A.7.27) $$g_A^2 k\{1+k(1-\alpha)\} - g_A(n+\theta\alpha) - n(n+\theta)(1-\alpha) = 0$$

(A.7.28) $$\hat{g}_A = \frac{n+\theta\alpha+\sqrt{(n+\theta\alpha)^2+4k\{1+k\{1+k(1-\alpha)\}n(n+\theta)(1-\alpha)\}}}{2k\{1+k(1-\alpha)\}}$$

In agreement with the steady state growth rate of per capita income $\gamma - n$ in (30), we find $\gamma - n = \frac{\hat{g}_A}{1-\alpha}$. The value of $\omega - g_A$ at the maximum of g_A is

(A.7.29) $$\omega - g_A\left(g_A = \frac{n+\theta}{k}\right) = \frac{n+\theta}{k}\frac{-\theta(1+k)}{\alpha k g_K+n\{1+2k-\alpha k\}+2\theta(1+k)}$$

At $g_K = \frac{g_A}{1-\alpha} + n = \frac{n+\theta}{k(1-\alpha)} + n$ this becomes

(A.7.30) $$\omega - g_A\left(g_A = \frac{n+\theta}{k};\ g_K = \frac{g_A}{1-\alpha} + n\right) = \frac{1}{k}\frac{-\theta(1+k)(1-\alpha)(n+\theta)}{(n+\theta)\{1+2k(1-\alpha)\}+\theta(1-\alpha)}$$

The value of $\omega - g_A$at the minimum of g_A is

(A.7.31) $$\omega - g_A\left(g_A = \frac{n}{k}\right) = \frac{\theta\{\alpha g_K+(1-\alpha)n\}}{\alpha k g_K+n\{1+2k-\alpha k\}+\theta(1+k)}$$

At $g_K = \frac{g_A}{1-\alpha} + n = \frac{n}{k(1-\alpha)} + n$ this becomes

(A.7.32) $$\omega - g_A\left(g_A = \frac{n}{k};\ g_K = \frac{g_A}{1-\alpha} + n\right) = \frac{\theta n\{\alpha+k(1-\alpha)\}}{kn\{1+2k(1-\alpha)\}+\theta(1+k)k(1-\alpha)}$$

The progression of minima and maxima explained in the main text is now easily derived. From some time t_0^* on, the lowest possible value of g_A is $\frac{n}{k}$. Denote this minimum as $g_{A,1}$ and let the moment from which this minimum is first achieved The corresponding attractor of g_K is $g_{K,1} = \frac{g_{Amin,1}}{1-\alpha} + n$. If g_A is at this minimum there is, irrespective of the initial value of g_K, a t_1 such that for all $t \geq t_1$ the value of g_K exceeds $g_{K,1} - \Delta$ ($\Delta > 0$), where Δ can be chosen arbitrarily small by selecting t_1 large enough. This implies $\omega > \omega\big(g_A = g_{Amin,1}, g_K = g_{K,1} - \Delta\big) = \omega_1$ for all $t \geq t_1$. Thus there is a $t_1^* \geq t_1$ such that (cf. figure A2) $g_A > \omega_1 - \Delta = g_{Amin,2} > g_{Amin,1}$ for all $t \geq t_1^*$.

Continuing the same process we obtain (the $\Delta's$ can be omitted because in the next step they become superfluous) a sequence $t_i^* \geq t_i \geq t_{i-1}^*, i = 1, 2, ...$ such that

- $g_A > g_{Amin,i}$ for $t \geq t_{i-1}^*$
- $g_K > g_{K,i} = \frac{g_{Amin,i}}{1-\alpha} + n$ for $t \geq t_i$
- $g_A > g_{Amin,i+1} = \omega_i = \omega\big(g_A = g_{Amin,i}, g_K = g_{K,i}\big)$ for $t \geq t_i^*$

Since $g_{Amin,i+1} > g_{Amin,i}$ as long as $\omega_i < \overline{\omega}$, the sequence $g_{Amin,i}$ converges to $\overline{\omega}\left(g_A = \hat{g}_A;\ g_{K,i} = \frac{\hat{g}_A}{1-\alpha} + n\right) = \hat{g}_A$. The reasoning for the maximum values of g_A and g_K, is analogous, *mutatis mutandis.*

Section 8

The solution for *Y* that can be computed from (47), (48) and (50) is

(A.8.1) $$Y = \left\{\frac{k}{k-1} L_Y{}^{1-\alpha} \left(\frac{s_K}{\gamma}\right)^{\alpha}\right\}^{\frac{(k^2\hat{g}_A+n+\theta)}{\sigma}} \left(\frac{1+k}{\hat{g}_x}\frac{k s_D}{\delta_c}\right)^{\frac{n+\theta-k\hat{g}_A}{\sigma}} \left(\chi^k e^{nt}\right)^{\frac{(k+1)\hat{g}_A}{\sigma}}$$

Here $$\sigma = -\alpha(n+\theta) + k\hat{g}_A\{k(1-\alpha)+1\}$$

$$\hat{g}_A = (\gamma - n)(1-\alpha)$$

$$\frac{\hat{g}_x}{1+k} = (\gamma - n)\{1 + k(1-\alpha)\} + n$$

$$\chi = B_0\left(\frac{s_B}{1-\alpha+s_B} L_0\right)^{\frac{1}{k}} \frac{k}{n+\theta}$$

The marginal consumption effect of capital outlays and the two categories of research expenditure can be obtained by differentiating *Y* with respect to the s_i first, next derive the quasi elasticities of *C* with respect to the s_i, which are interesting in their own right, and finally replace the changes in e ds_i by the appropriate expressions in dU_i.

Thus the first step is to differentiate (A.8.1) with respect to s_K, s_D, L_Y, and χ.

(A.8.2) $$\frac{dY}{Y} = \frac{n+\theta+k^2\hat{g}_A}{\sigma}\left\{(1-\alpha)\frac{dL_Y}{L_Y} + \alpha\frac{ds_K}{s_K}\right\} + \frac{n+\theta-k\hat{g}_A}{\sigma}\frac{ds_D}{s_D} + \frac{(1+k)k\hat{g}_A}{\sigma}\frac{d\chi}{\chi}$$

Express $d\chi$ and dL_Y in ds_B (use 45 in case of L_Y):

(A.8.3) $$\frac{d\chi}{\chi} = \frac{1}{k}\frac{1-\alpha}{1-\alpha+s_B}\frac{ds_B}{s_B}; \ \frac{dL_Y}{L_Y} = \frac{-s_B}{1-\alpha+s_B}\frac{ds_B}{s_B}$$

Consequently:

(A.8.4) $$\frac{dY}{Y} = \frac{n+\theta+k^2\hat{g}_A}{\sigma}\alpha\frac{ds_K}{s_K} + \frac{n+\theta-k\hat{g}_A}{\sigma}\frac{ds_D}{s_D} + \frac{ds_B}{s_B}\frac{(1-\alpha)}{\sigma(1-\alpha+s_B)}[(1+k)\hat{g}_A - s_B\{n+\theta+k^2\hat{g}_A\}]$$

Next use (46) to obtain dC:

(A.8.5) $$\frac{dC}{C} = \frac{dY}{Y} - \frac{s_D}{1-s_K-s_D}\frac{ds_D}{s_D} - \frac{s_K}{1-s_K-s_D}\frac{ds_K}{s_K} = \left[\frac{\alpha\{n+\theta+k^2\hat{g}_A\}}{\sigma s_K} - \frac{1}{1-s_K-s_D}\right]ds_K + \left[\frac{\{n+\theta-k\hat{g}_A\}}{\sigma s_D} - \frac{1}{1-s_K-s_D}\right]ds_D + ds_B\frac{(1-\alpha)}{\sigma(1-\alpha+s_B)}\left[\frac{(1+k)\hat{g}_A}{s_B} - \{n+\theta+k^2\hat{g}_A\}\right]$$

The quasi elasticities in (51) are taken from (A.8.5). To obtain the marginal consumption effects in (51) first write the s_i in terms of C:

(A.8.6) $$s_i = \frac{U_i}{Y} = \frac{U_i}{C+U_K+U_D}$$

Differentiate:

(A.8.7) $$ds_i = \frac{(C+U_K+U_D)dU_i - U_i(dC+dU_K+dU_D)}{(C+U_K+U_D)^2} = (1-s_K-s_D)\left[\frac{dU_i}{C} - s_i\left(\frac{dC}{C}+\frac{dU_K}{C}+\frac{dU_D}{C}\right)\right]$$

A change in budget U_i does not alter the s_j $(j \neq i)$, implying

(A.8.8) $$\frac{dU_j}{C} - s_j\left(\frac{dC}{C}+\frac{dU_K}{C}+\frac{dU_D}{C}\right) = 0 \ \ (j \neq i)$$

For $ds_K \neq 0$ equation (A.8.8) leads to:

(A.8.9) $$\frac{dU_D}{C} = \frac{s_D}{1-s_D}\left(\frac{dC}{C}+\frac{dU_K}{C}\right)$$

Substitute this into (A.8.7) and obtain:

(A.8.10) $$ds_K = \frac{1-s_K-s_D}{1-s_D}\left[(1-s_D-s_K)\frac{dU_K}{C} - s_K\frac{dC}{C}\right]$$

Substitute this result in the quasi elasticity with respect to s_K in (A.8.5) to obtain (51a). The result for U_D is symmetric. In case of U_B, (A.8.8) leads to $dU_j = \frac{s_j}{1-s_K-s_D}dC$ $(j = K, D)$. Substitute this into (A.8.7) and the result in the quasi elasticity of C with respect to s_B in order to obtain (51c).

The golden rule values for s_K and s_D need to be determined simultaneously, since their values depend on each other. Setting both quasi-elasticities of C equal to zero we obtain two equations:

(A.8.11) $$\alpha\{n+\theta+k^2\hat{g}_A\} = \frac{\sigma s_K}{1-s_K-s_D}$$

(A.8.12) $$\{n+\theta-k\hat{g}_A\} = \frac{\sigma s_D}{1-s_K-s_D}$$

The golden rule values are easily derived by solving s_K and s_D from these two equations; the golden rule value of s_B the relevant quasi elasticity equal to zero.