POPULATION AND IDEAS: A THEORY OF ENDOGENOUS GROWTH

Charles I. Jones

Working Paper **6285**

NBER WORKING PAPER SERIES

POPULATION AND IDEAS: A THEORY OF ENDOGENOUS GROWTH

Charles I. Jones

Working Paper 6285
http://www.nber.org/papers/w6285

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1997

I would like to thank Robert Barro, Robert Lucas, Paul Romer, T.N. Srinivasan, John Williams, and Alwyn Young for insightful comments. The Center for Economic Policy Research at Stanford, the Hoover Institution, and the National Science Foundation (SBR-9510916) provided financial support. Copies of this paper may be downloaded from the world wide web at http://www.stanford.edu/~chadj/index.html. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

Population and Ideas: A Theory of Endogenous
Growth
Charles I. Jones
NBER Working Paper No. 6285
November 1997
JEL Nos. O41, O30

ABSTRACT

Why do economies exhibit sustained growth in per capita income? This paper argues that endogenous fertility and increasing returns to scale are the fundamental ingredients in understanding endogenous growth. Endogenous fertility leads the scale of the economy to grow over time. Increasing returns translates this increase in scale into rising per capita income. A justification for increasing returns rather than linearity in the equation for technological progress is the fundamental insight of the idea-based growth literature according to this view. Endogenous fertility together with the increasing returns associated with the nonrivalry of ideas generates endogenous growth.

Charles I. Jones
Department of Economics
Stanford University
Stanford, CA 94305-6072
and NBER
Chad.Jones@Stanford.edu

1 Introduction

Why do economies exhibit sustained per capita growth? Almost by definition, a model of long-run growth requires a differential equation that is "linear" in its state variable, such as

$$\dot{X} = \_\_X. \tag{1}$$

Growth models differ according to the way in which they label the X variable and the story they tell in order to fill in the blank.[1] For example, the original Solow model without technological progress does not have linearity because of diminishing returns to capital, so that growth eventually ceases. To generate sustained growth in that model, one adds exogenous technological progress in the form of a differential equation that is assumed to be linear: $\dot{A} = gA$.

Much of the work in both new and old growth theory can be read as the search for the appropriate characterization of equation (1). Early work with "AK" models sought linearity by eliminating the diminishing returns to capital accumulation, either in the form of physical capital or with human capital, or with both.[2] Later work by Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992) returned the linearity to the equation for technological progress and focused on theories in which intentional research effort by profit-maximizing firms fills the blank in equation (1).

While the proverbial jury is still out, a strong case can be made that this search has so far been unsuccessful. First, there is generally very little empirical support for the first generation AK models.[3] The case against the

[1]This way of summarizing growth models is taken from Romer (1995).

[2]See the well-known work of Romer (1987), Lucas (1988), and Rebelo (1991).

[3]Several types of evidence are relevant. First, cross-country growth regressions such as Mankiw, Romer and Weil (1992) uniformly find evidence against linearity, suggesting that the returns to capital are at most 0.7 or 0.8. Second, many AK models rely on learning-by-doing or externalities. Micro evidence in Lucas (1993) suggests that learning-by-doing within a product line is bounded, and the literature has failed to produce any evidence of

second generation "idea-based" growth models of Romer and others is even more compelling. These models suffer from the well-known problem that the growth rate is increasing in the size of the economy and should explode in the presence of population growth.[4]

In this paper, however, I'd like to draw attention to a more fundamental problem with existing theories of long-run growth. Exact linearity in these models — and in any endogenous growth model — is critical: if the model is slightly more than linear then growth explodes, while if the model is slightly less than linear then there is no long-run growth.[5] Since exact linearity is obviously so important, any plausible explanation of long-run growth should possess an intuitive and compelling justification for linearity. On this basis, existing models are clearly deficient. The linearity in existing models is assumed ad hoc, with no motivation other than that we must have linearity somewhere to generate endogenous growth. For this reason, existing explanations of long-run growth are inadequate.

In this paper, I develop a new theory of endogenous growth in which linearity is motivated from first principles. The first key ingredient of this model is endogenous fertility. At an intuitive level, the reason why endogenous fertility helps is straightforward. Consider a standard Solow-Swan model. With the labor force as a factor that cannot be accumulated endogenously, one has to look for a way — typically arbitrary — to eliminate the diminishing returns to physical capital. In contrast, with an endogenously accumulated labor force, both capital and labor are accumulable factors, and a standard constant returns to scale setup can easily generate an endogenously growing economy.

However, endogenous fertility in a model with constant returns to scale

large externalities. Finally, time series evidence like that in Jones (1995b) generally fails to support the implications of AK models: investment rates in physical or human capital have risen in a number of countries with no corresponding increase in long-run growth.

[4]See Jones (1995a).

[5]The knife-edge nature of linearity has been noted by Solow (1994), among others.

in all production functions will not generate endogenous growth in *per capita* variables. This leads to the second key ingredient of the model, increasing returns to scale. Endogenous fertility leads to endogenous growth in the scale of the economy. Increasing returns to scale in the production function for aggregate output translates the endogenous growth in scale into endogenous growth in per capita output.

The fertility side of this paper builds on the endogenous fertility literature associated with Dasgupta (1969), Pitchford (1972), Razin and Ben-Zion (1975), and Becker and Barro (1988) and provides the compelling justification for linearity absent in previous models. Individuals choose to have a certain number of children, taking into account the costs — such as hospital bills and foregone labor income — and benefits — such as the utility value of offspring — of fertility. The number of children per adult, which we'll call $\tilde{n}$, will (by definition) be constant in steady state. Suppose people live for one discrete period. With N people in the population, the total number of offspring is equal to the number of children per adult multiplied by the size of the population: $N_{t+1} = \tilde{n}N_t$. But then, of course, the net increase in population is given by

$$N_{t+1} - N_t = nN_t,$$

where $n \equiv \tilde{n} - 1$. Or, in continous time, $\dot{N} = nN$. By choosing the number of children to have, individuals choose the *proportional* rate of increase of the population. The law of motion for population exhibits linearity because people reproduce in proportion to their number. This is very different from the production process for physical capital or human capital or even ideas. There is no reason why, for example, individuals who spend a constant fraction of their time endowment accumulating skills should increase their human capital by a constant proportion over time. And there is no reason why a constant research effort should increase the stock of ideas or the level of productivity by a constant proportion every period. But it is a biological fact of nature that people reproduce in proportion to their number.[6]

Research on idea-based growth models provides a justification for increasing returns, the second key ingredient of the model, that is based on first principles. At least since Phelps (1966) and Shell (1967), economists have recognized that the nonrivalry of knowledge implies that aggregate production is characterized by increasing returns to scale. This argument is made very clearly by Romer (1990). Ideas are nonrivalrous; they can be used at any scale of production after being produced only once. For example, consider the production of any new product, say the digital videodisc player or the latest world wide web browser. Producing the very first unit may require considerable resources — the product must be invented or designed. However, once the product is invented, it never needs to be invented again, and the standard replication argument implies that subsequent production occurs with constant returns to scale. Including the production of the "idea," or the design of the product, production is characterized by increasing returns. This property, rather than the assumption that the differential equation governing technological progress is linear, is the key contribution we need from the idea-based growth literature.[7]

According to this paper, the story underlying sustained, long-run per capita growth is more complicated than previously thought. The inherent nonrivalry of ideas means that the economy is characterized by increasing returns to scale. Economic growth occurs because the economy is repeatedly discovering newer and better ways to transform labor into valuable goods

[6]It is important to note that the linearity of the production technology for offspring does not guarantee positive growth. For example, people may, subject to their economic environment, choose to have one child per adult in steady state, generating a stable population. This can be the case if the production technology for output exhibits decreasing returns to scale due to a fixed factor such as land. Linearity is not sufficient for growth, but it is necessary.

[7]Alternative methods for introducing increasing returns to scale in the model, such as a Marshallian externality associated with capital accumulation, will also lead to endogenous growth. I focus on the idea-based theory of increasing returns because it can be motivated from first principles.

and services. However, the creation of new ideas by itself is not sufficient to generate sustained growth. For example, suppose an economy invents 100 new ideas every year. As a fraction of the (ever evolving) existing stock of ideas, these 100 new ideas become smaller and smaller. Sustained growth requires that the number of new ideas itself grow exponentially over time. This in turn requires that the number of inventors of new ideas grows over time, which requires population growth. Endogenous fertility, combined with the increasing returns to scale associated with ideas, delivers sustained long-run growth.[8]

This paper builds on a number of earlier insights. Jones (1995a) modified the Romer (1990) model to eliminate the counterfactual prediction that the growth rate of the economy depends on the size of the population. In the modified model, the growth rate of the economy depends on the growth rate of the population. In Jones (1995a), however, the population growth rate was assumed to be exogenously given, and it turned out that the long-run growth rate of the economy was invariant to policy changes. Here, the population growth rate is endogenized, and policy changes can affect the long-run growth rate of the economy through their effects on fertility. Because the channel through which policy affects growth is fertility, however, the nature of the effects of policy on long-run growth is often counter to conventional wisdom. For example, subsidies to R&D and capital accumulation, even though they may be welfare improving, will reduce long-run growth in the model.

Kremer (1993) is also closely related and provides the most compelling supporting evidence for the model developed here. This evidence will be reviewed later in the paper. Young (1995) is related in two ways. First, it

[8]One of the cases considered by Raut and Srinivasan (1994) also delivers this result. They consider a model of endogenous fertility in which the productivity level of firms depends, through an externality, on the stock of labor in the economy. The emphasis in their paper, however, is that the model can exhibit faster than exponential growth and even chaotic dynamics depending on the nature of the external effect.

represents an alternative theory of economic growth that does not appear to be inconsistent with any well-accepted empirical evidence. With no population growth, positive per capita growth in the long run in Young's model exists because of quality upgrading, but the long-run growth rate is invariant to policy. Second, if one adds population growth to Young's model, per capita growth is increased by an additive term that is proportional to population growth. In this sense, the properties of the endogenous fertility model developed here are directly relevant to Young's model.

Section 2 of this paper develops the decentralized model of endogenous growth in the context of "basic science." That is, the model is based on the assumption that the ideas underlying growth are not only nonrivalrous, they are pure public goods. This assumption is employed because it simplifies the analysis considerably, but it may also be of independent interest. First, the fundamental discoveries of basic research may be an important driving force underlying growth. Second, it is sometimes conjectured that basic science should be modeled as an exogenous process, like exogenous technical progress in a Solow model. The analysis here suggests that insight is gained by moving beyond this view. For example, even if the ideas of basic science fall from above like apples from trees, the fertility channel and increasing returns are crucial: the number of Isaac Newtons depends on the the size of the population that is available to sit under trees.

Section 3 explores the welfare properties of the model. Section 4 contains a general discussion of the model's predictions, and Section 5 concludes.

2 The Decentralized Model

Models of endogenous fertility are typically somewhat difficult to solve. In what follows, I have chosen a particular theory and made particular assumptions to minimize the effort required to get to the basic results. I will indicate in the appropriate places how the results generalize.

2.1 Preferences

One of the key insights of Barro (1974) was to think about utility-maximizing individuals who care not only about their own consumption but also about their children's consumption. This reasoning was extended by Razin and Ben-Zion (1975) and Becker and Barro (1988) to model endogenous fertility: parents also care about the number of children that they have, and there may be costs to increasing the number of offspring.

The time s utility of a representative agent born at time 0 is given by

$$U_{0,s} = \int_s^\infty e^{-\rho(t-s)} u(c_t, \tilde{N}_{0,t})\, dt, \tag{2}$$

Individuals live for only an instant, during which time they work, consume, and reproduce. However, they also leave bequests and care about the consumption and number of their offspring. This is a continuous time version of a dynastic utility function like that considered by Becker and Barro (1988). c_t is the consumption of a representative member of the population at time t, N_t is the number of people in generation t, and $\rho > 0$ is the rate of time preference that applies across generations. $\tilde{N}_{0,t} \equiv N_t/N_0$ represents the number of offspring of a member of generation 0 that are alive at time t.

With respect to the kernel of the utility function, it turns out to be convenient to assume

$$u(c, \tilde{N}) = \log c + \epsilon \log \tilde{N}, \tag{3}$$

where $\epsilon > 0$. Both the marginal utility of consumption and the marginal utility of progeny are positive but diminishing. The elasticity of substitution between consumption and progeny is one, as in Barro and Becker (1989). Within the class of utility functions with a constant elasticity of substitution between consumption and progeny, this unit elasticity is required for the existence of a balanced growth path, as we will see shortly.[9] It also

[9]This restriction is closely related to the restriction in dynamic general equilibrium

guarantees that the dynastic approach is time consistent — choices made by the dynastic head of generation 0 will be implemented by subsequent generations.

Finally, characterizing the equilibrium of the model is much easier under the stronger assumption that $\epsilon = 1$, so that per capita consumption and offspring receive equal weights in the utility function. In the presentation of the model, we will make this assumption and indicate at the appropriate time what happens when $\epsilon \neq 1$.

2.2 Technology

The consumption-capital good in the economy, final output Y, is produced according to

$$Y = A^{\sigma} K^{\alpha} L_Y^{1-\alpha}, \tag{4}$$

where A is the stock of ideas in the economy, K is capital, L_Y is labor, and the parameters satisfy $\sigma > 0$ and $0 < \alpha < 1$. While this kind of production function is commonly used in economics, it incorporates a fundamental insight into the process of economic growth. Specifically, the production function exhibits increasing returns to scale because of the nonrivalry of ideas. Nonrivalry means that ideas only need to be discovered once: after an idea is discovered, it can be used at any scale of production. For example, the production process for light bulbs presumably is well-characterized today by constant returns to scale: doubling the number of factories, workers, and materials will double the number of light bulbs produced. However, this was not always the case. Edison expended a great deal of perspiration and inspiration to produce the first light bulb. But once the design was perfected, the same amount of labor and factory time that was needed to produce the first light bulb could presumably produce thousands of light

business cycle models that consumption must enter in log form if consumption and leisure are additively separable (leisure per person does not need to enter in log form because it is not growing over time).

bulbs. That is, including the effort required to produce the idea underlying the light bulb, production is characterized by increasing returns.

Similarly, holding the stock of ideas A constant, the production function in equation (4) exhibits constant returns to capital and labor, according to the usual replication argument. However, production exhibits increasing returns once the stock of ideas is taken into account. The strength of increasing returns is measured by σ.

The technology for producing offspring is straightforward. Individuals are endowed with one unit of labor, and generating a fertility rate of n requires $\beta(n)$ units of time, with $\beta'(n) > 0$. The time that individuals have left over to supply to the labor market is $1-\beta(n)$. We assume that $\beta(0) = 0$ and $\lim_{n\longrightarrow\infty} \beta(n) > 1$ so that there is an upper bound to the fertility rate of an individual. To generate an interior solution in the decentralized model, we also require $\beta''(n) > 0$, but this assumption can be relaxed in alternative formulations of fertility theory.

Reproduction occurs through asexual budding; there is no distinction between male and female agents. With N identical agents in the economy, the total number of offspring produced in an economy with individual fertility n is given by

$$\dot{N} = nN. \tag{5}$$

Capital accumulates in this economy in the form of assets owned by households. Letting v denote the per capita stock of assets ($K = Nv$ is imposed later),

$$\dot{v} = (r-n)v + w(1-\beta(n)) - c - f, \tag{6}$$

where r is the market return on assets, w is the wage rate per unit of labor, and f represents per capita lump sum taxes collected by the government ($f \equiv F/N$).

The final component of the technology of the economy is the production of ideas. New ideas are produced by researchers according to

$$\dot{A} = \bar{\delta} L_A, \tag{7}$$

where L_A denotes labor engaged in research, and $\dot{A}$ represents the measure of new ideas created at a point in time. The resource constraint on labor is

$$L_A + L_Y = (1 - \beta(n))N \equiv L. \tag{8}$$

While individual researchers who are small relative to the total number of researchers take $\bar{\delta}$ as given, in fact it may depend on features of the aggregate economy. The true relationship between new ideas and research is assumed to be given by

$$\dot{A} = \delta L_A^{\lambda} A^{\phi}, \tag{9}$$

where $\delta > 0$, $0 < \lambda \leq 1$ and $\phi < 1$ are parameters. This formulation allows for both positive and negative externalities in research. At a point in time, congestion or duplication in research may reduce the social value of a marginal unit of research, associated with $\lambda < 1$. In addition, the productivity of research today may depend on the stock of ideas discovered in the past. The case $\phi > 0$ corresponds to a situation in which the productivity of research increases with past discoveries (knowledge spillovers). The case of $\phi < 0$ suggests that research gets harder as more ideas are discovered. Finally, with $\phi = 0$, these two effects offset and the productivity of research is independent of the number of ideas discovered in the past. Note that equation (9) allows for increasing, constant, or decreasing returns to scale in the production of new ideas.

2.3 Market Structure

Romer (1990) and others have emphasized that ideas are nonrivalrous but partially excludable. The assumption that ideas are at least partially excludable allows inventors to capture some of the social value that they create.

This feature, together with the increasing returns to scale implied by nonrivalry, leads Romer, Grossman and Helpman, and Aghion and Howitt to favor models with profit-maximizing entrepreneurs and imperfect competition — what we might call "Silicon Valley" models.

Here, we will make an alternative assumption which will have the flavor of growth through basic science. In particular, we assume that ideas are nonrivalrous and nonexcludable; that is, they are pure public goods. This means that inventors cannot use the market mechanism to capture any of the social value they create. In the absence of some non-market intervention, no one would become a researcher because of the fundamental ineffectiveness of property rights over basic science, and no growth would occur.

This alternative assumption serves two purposes. First, in terms of the modeling, this assumption greatly simplifies the analysis of the decentralized model. We assume that all markets are perfectly competitive, and then introduce a government to collect lump-sum taxes and use the revenues to fund research publicly. Second, this case may be of independent interest. Basic research is to a large extent publicly-funded and can be characterized as a pure public good.[10] Previous "Silicon Valley" style models have analyzed the case in which research is undertaken by private entrepreneurs who are compensated through imperfectly competitive markets. This paper explores the alternative extreme in which growth is associated with basic science undertaken by publicly-funded scientists.[11]

The government collects lump-sum taxes F from individuals and uses this revenue to hire research scientists at the market wage w. We assume that the government collects as much revenue as needed so that a constant fraction of the labor force, $0 < \bar{s} < 1$, is hired as researchers: i.e. $L_A = \bar{s}L$.

[10]Or at least approximately so. Successful researchers do seem to capture some fraction of the value they create through prestige.

[11]See Shell (1967) for an early application of this approach.

2.4 Equilibrium

A competitive equilibrium in this model is a collection of quantities $\{c_t,\ Y_t,\ K_t,\ A_t,\ v_t,\ L_{Yt},\ L_{At},\ N_t,\ n_t\}$, prices $\{w_t, r_t\}$, and lump-sum taxes $\{F_t\}$ such that

- Individuals choose $\{c_t, n_t\}$ to maximize dynastic utility in equation (2) subject to the laws of motion for asset accumulation (6) and population (5), taking $\{r_t, w_t, F_t\}$ as given.
- Firms producing output rent capital K and labor L_Y to maximize profits, taking the rental prices r and w and the stock of ideas A as given.
- Markets clear at the prices $\{w_t, r_t\}$ and the taxes $\{F_t\}$. In particular, the stock of assets held by consumers V is equal to the total capital stock K, and the number of researchers is a constant fraction $\bar{s}$ of the labor force.

We now characterize the competitive equilibrium along a balanced growth path, i.e. when all variables are growing at constant (exponential) rates.

The first-order conditions from the utility maximization problem for individuals, evaluated along a balanced growth path, imply that the fertility rate chosen by the household satisfies:

$$\frac{(r - g_Y)(v + w\beta'(n))}{\tilde{N}} = \frac{u_{\tilde{N}}}{u_c}. \tag{10}$$

This equation is the dynamic equivalent of the condition that the marginal rate of transformation (the left-hand side) equals the marginal rate of substitution (the right-hand side) between people and consumption. The marginal rate of transformation is based on the cost to the individual of increasing fertility, which involves two terms. First, there is a capital-narrowing effect: adding to the population dilutes the stock of assets per person. Second, there is the direct cost of wages that are foregone in order to increase the population growth rate. The total cost is scaled by the size of the population so that it is measured in terms of people rather than as a rate, and it is multiplied by the effective discount rate $r - g_Y$ to put it on a flow basis. This marginal rate of transformation is equal to the static marginal rate of substitution $u_{\tilde{N}}/u_c$ along the optimal balanced growth path.

This relationship makes it clear why a unit elasticity of substitution between people and consumption is required. The marginal rate of transformation on the left-hand side of equation (10) will end up being proportional to $y/\tilde{N}$, where y is per capita output, Y/N. Therefore, the marginal rate of substitution must be proportional to $c/\tilde{N}$ for a balanced growth path to exist; otherwise, the cost and the benefit of fertility will grow at different rates and the economy will be pushed to a corner. The equation also makes clear why the curvature $\beta''(n) > 0$ is required: with $\beta(n) = 1 - \beta n$, for example, equation (10) doesn't directly depend on n, and households will move to a corner solution.

Other first-order conditions characterizing the balanced growth path equilibrium are more familiar. For example, consumption growth satisfies the following Euler equation:

$$\frac{\dot{c}}{c} = r - n - \rho. \tag{11}$$

Also, the first-order conditions from the firm's profit-maximization problem, assuming no depreciation, are

$$r = \alpha Y/K$$

and

$$w = (1-\alpha)Y/L_Y = (1-\alpha)y \cdot \frac{1}{1-\beta(n)} \cdot \frac{1}{1-\bar{s}}. \tag{12}$$

With these first-order conditions in mind, we are ready to characterize the steady-state growth rate of the economy. Along the balanced growth path, the key growth rates of the model are all given by the growth rate of the stock of ideas:

$$g_y = g_k = g_c = \frac{\sigma}{1-\alpha} g_A, \tag{13}$$

where g_z denotes growth rate of some variable z along the balanced growth path, y is per capita income Y/N, and k is capital per person K/N.[12]

The growth rate of ideas, g_A, is found by dividing both sides of equation (9) by A:

$$\frac{\dot{A}}{A} = \delta \frac{L_A^{\lambda}}{A^{1-\phi}}.$$

Along a balanced growth path, the numerator and the denominator of the right-hand side of this expression must grow at the same rate, and this requirement pins down the growth rate of A as

$$g_A = \frac{\lambda}{1-\phi} g_{L_A}.$$

Finally, along a balanced growth path, L_A must grow at the rate of growth of the population. Therefore,

$$g_A = \frac{\lambda n}{1-\phi}. \tag{14}$$

Combining this result with equation (13),

$$g_y = \gamma n, \tag{15}$$

where $\gamma \equiv \frac{\sigma}{1-\alpha}\frac{\lambda}{1-\phi}$.

As in Jones (1995a), the per capita growth rate of the economy is proportional to the population growth rate.[13] This is a direct consequence of

[12]This relationship is derived as follows. First, the constancy of consumption growth requires a constant interest rate and therefore a constant capital-output ratio, yielding the first equality. Second, the asset accumulation equation in (6) is simply a standard capital accumulation equation. For the capital stock to grow at a constant rate, the capital-consumption ratio must be constant, yielding the second equality. Finally, log-differentiating the production function in (4) yields the last equality.

[13]Phelps (1966), Nordhaus (1969), and Judd (1985) are early idea-based models that contain the result that the per capita growth rate of the economy is proportional to the population growth rate. Arrow (1962) and Sheshinski (1967) generate this result in models in which the increasing returns to scale is due to external learning by doing. All of these models, however, take the population growth rate as exogenous.

increasing returns to scale: with $\sigma = 0$, there is no per capita growth in the long-run. Notice that balanced growth in the presence of population growth in this model requires $\alpha < 1$ and $\phi < 1$. That is, the capital accumulation equation and the law of motion for ideas must both be less than linear in their own state variables; otherwise, growth explodes and the level of consumption and income is infinite in a finite amount of time.

2.5 Fertility in the Decentralized Economy

The rate of population growth is determined by consumer optimization, as in equation (10). Using the fact that $\epsilon = 1$ and $r - g_Y = \rho$ along a balanced growth path, and substituting for the wage from equation (12), equation (10) can be written as

$$k + (1-\alpha)y \cdot \frac{\beta'(n)}{1-\beta(n)} \cdot \frac{1}{1-\bar{s}} = \frac{1}{\rho}c. \tag{16}$$

Some algebra then shows that along a balanced growth path, the rate of fertility satisfies[14]

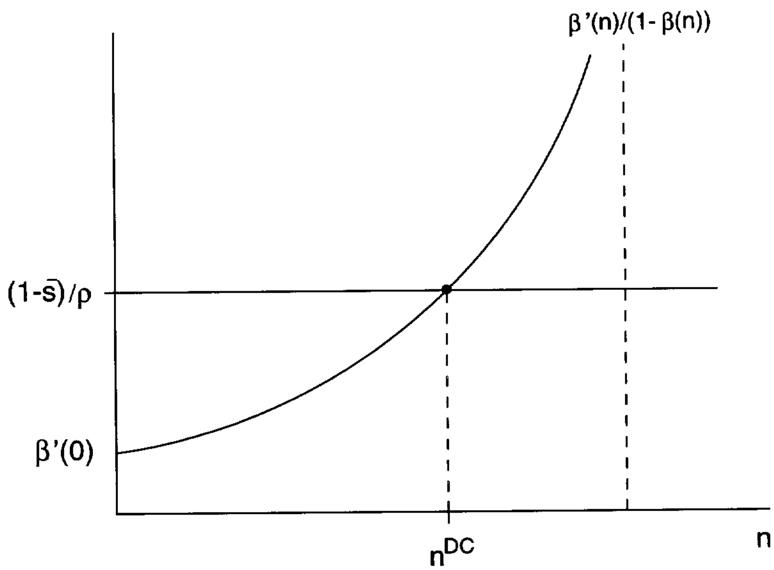
$$\frac{\beta'(n^{DC})}{1-\beta(n^{DC})} = \frac{1-\bar{s}}{\rho}. \tag{17}$$

The solution to this equation exists and is unique under the assumption that $\beta'(0) < (1-\bar{s})/\rho$, as shown in Figure 1. Recall that the relationship in equation (15) that $g_y = \gamma n$ then determines the growth rate of the economy along the balanced growth path.

The steady state growth rate of the economy is directly proportional to the fertility rate. This rate is smaller the higher is the rate of time preference ρ or the higher is the cost of fertility $\beta(\cdot)$. Interestingly, the growth rate of the economy is *decreasing* rather than increasing in the fraction of the labor force devoted to research. This is very different from the results in previous idea-based growth models and reflects the fact that growth is driven by

[14]Specifically, divide both sides of the equation by k and use the fact that $y/k = r/\alpha$ and $c/k = y/k - g_Y = (1-\alpha)/\alpha * r + \rho$ along a balanced growth path.

Figure 1: Solving for n^{DC}



a different mechanism. Here, changes in research intensity affect long-run growth only through their effect on fertility. A larger research sector raises the marginal product of labor in the output sector and therefore raises the wage. This means that the opportunity cost of fertility is higher, which reduces population growth and therefore reduces steady-state per capita growth.

3 Welfare and a Planner Problem

With more than one generation of agents, it is not obvious how to define social welfare: it depends on how one weights the utility of different generations. We focus on a narrower question: does the allocation of resources achieved in the market economy maximize the utility of each generation given the initial conditions that constrain their choices?

To maximize the welfare of a representative generation (the generation alive at time zero here), the social planner solves

$$\max_{\{c,s,n\}} U_0 = \int_0^\infty e^{-\rho t} u(c_t, \tilde{N}_{0,t})\ dt, \tag{18}$$

subject to

$$\dot{k} = A^\sigma k^\alpha (1-s)^{1-\alpha}(1-\beta(n))^{1-\alpha} - c - nk, \tag{19}$$

$$\dot{A} = \delta s^\lambda (1-\beta(n))^\lambda N^\lambda A^\phi, \tag{20}$$

and

$$\dot{N} = nN. \tag{21}$$

The first order conditions from this maximization problem can be combined to yield several equations of interest. First, optimal consumption satisfies a standard Euler equation

$$\frac{\dot{c}}{c} = \alpha\frac{y}{k} - n - \rho. \tag{22}$$

Second, the first order conditions together with the equations governing the law of motion for capital and ideas can be solved along a balanced growth path to yield optimal research intensity:

$$s^{SP} = \frac{1}{1+\psi^{SP}}, \tag{23}$$

where

$$\psi^{SP} = \frac{1-\alpha}{\lambda\sigma}\left(\frac{\rho(1-\phi)}{\lambda n} + 1 - \phi\right).$$

To solve for the steady state rate of population growth, we follow the steps used for the decentralized model. The first order conditions from the planner's problem can be combined to yield a condition analogous to equation (10):

$$\left(k + (1-\alpha)y \cdot \frac{\beta'(n)}{1-\beta(n)} \cdot \frac{1}{1-s^{SP}}\right)\frac{\rho}{\tilde{N}} = \frac{u_{\tilde{N}}}{u_c} + \mu_2 \lambda \frac{\dot{A}}{\tilde{N}}, \tag{24}$$

where μ_2 is the shadow value of an idea (the co-state variable corresponding to equation (20)).

The distortion that affects fertility choice can be seen by comparing this equation to the corresponding condition in the decentralized model, either equation (10) or (16). Individual agents ignore the extra benefit associated with increasing returns to scale provided by additional population. This distortion is reflected by the presence of the second term on the right-hand side of equation (24), which corresponds to the utility value of the extra ideas created by an additional person.

Some additional algebra reveals that, along the balanced growth path, the optimal fertility rate satisfies[15]

$$\frac{\beta'(n^{SP})}{1-\beta(n^{SP})} = \frac{1}{\rho}. \tag{25}$$

Finally, the optimal steady-state growth rate of per capita income is given by $g_y^{SP} = \gamma n^{SP}$.

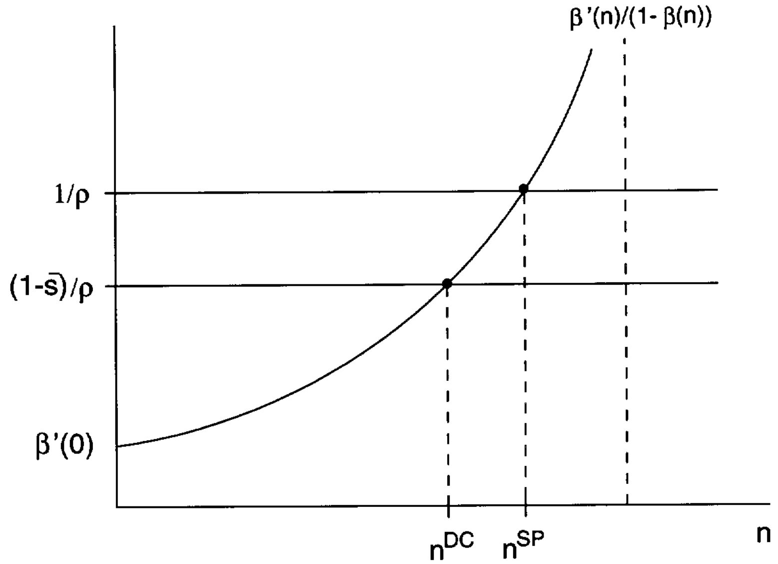
A comparison of equations (17) and (25) indicates that steady-state fertility and growth are inefficiently too slow in the decentralized economy, as shown in Figure 2. This results from the fact that, as noted above, individuals ignore the economy-wide benefit of fertility that is associated with increasing returns to scale: a larger population generates more ideas that benefit all agents in the economy.

In more general models that I have explored, this result can be overturned. For example, when the kernel of the utility function is generalized to place a higher weight on offspring, i.e. when $\epsilon > 1$, it is possible for the decentralized economy to have a fertility rate and therefore a growth rate that is inefficiently too high. This occurs if $\bar{s}$ is sufficiently smaller than s^{SP}.[16] Second, fertility and growth can be inefficiently high in a model in which the Romer (1990) market structure is used instead of the perfectly

[15]Once again, divide both sides of the equation by k and use the fact that $y/k = r/\alpha$ and $c/k = y/k - g_Y = (1-\alpha)/\alpha * r + \rho$ along a balanced growth path.

[16]To see part of the intuition, recall that from the standpoint of the decentralized economy, a lower research intensity increases fertility.

Figure 2: Comparing n^{DC} and n^{SP}



competitive/basic science market structure. With the imperfectly competitive market structure of Romer, capital is underpaid relative to its marginal product so that some resources are available to compensate entrepreneurs. However, recall that part of the opportunity cost of fertility is the additional capital that must be provided to offspring. Imperfect competition reduces this cost and can lead to inefficiently high fertility and growth.

4 Discussion

The model contains many novel predictions about the source of long-run growth. In particular, the "scale effects" prediction that has been a key problem in many endogenous growth models turns out to be a key feature in this model. Increasing returns to scale implies that the scale of the economy will matter. Instead of affecting (counterfactually) the long-run growth rate, however, scale affects the long-run *level* of per capita income. Large populations generate more ideas than small populations, and because ideas are nonrivalrous, the larger number of ideas translates into higher per capita income. Endogenous growth in the scale of the economy through fertility causes endogenous growth in per capita income.

Changes in government policies can affect the long-run growth rate by affecting the rate of fertility. For example, suppose that for each child, parents have to pay a fraction of their wages in taxes. Such a tax will reduce fertility and therefore reduce per capita growth.

Other policies can also affect population growth and per capita growth in the model, but the effects are often counterintuitive on the surface. Specifically, the imposition of many taxes in the model will increase rather than decrease growth. For example, a tax on labor income creates a wedge between working and child-rearing, the untaxed activity, and will increase fertility and per capita growth. A tax on capital reduces the opportunity cost of fertility by reducing the capital stock and wages and therefore will also increase growth. Finally, as we have already seen, an increase in an existing government subsidy to research will reduce long-run growth in the model. Notice that increasing the research subsidy may easily be welfare improving here, but not, as is often argued, because it increases the long-run growth rate.

What policies should the government follow in this model to obtain the socially optimal allocation of resources? It turns out that the policy is very simple and can be implemented without the use of lump-sum taxes. Suppose the government taxes labor income at rate τ_L and uses the revenue to fund research. In this case, it is easy to show that steady-state fertility in the decentralized economy satisfies

$$\frac{\beta'(n^{DC})}{1-\beta(n^{DC})} = \frac{1-\bar{s}}{1-\tau_L} \cdot \frac{1}{\rho}.$$

Moreover, the share of labor employed in research, $\bar{s}$, is equal to the tax rate τ_L. Therefore, by choosing a labor income tax rate of $\tau_L = s^{SP}$, both the fraction of labor working in research and the steady-state fertility and growth rate match the social optimum.[17]

How seriously should we take these general policy predictions? The clear prediction of the model, evident from equation (15), is that policies affect long-run growth only through their effects on fertility. This means that the standard intuitions taken from many other endogenous growth models do not apply. The intuition in this model is that the key wedge is between child-rearing and work in the market economy. Any policy that changes the relative price of these activities will affect fertility, and it is easy to see that the effects need not work in the conventional directions.

How should we test this model to judge its success? First, notice that, as with many idea-based growth models, this is a model of growth for the world economy as a whole. Therefore, testing the model as it stands with cross-sectional evidence is difficult. The Belgian economy does not grow solely or even primarily because of ideas invented by Belgians, so that the model does not predict that Belgium's per capita growth rate should be related to its population growth rate. Evidence that population growth rates and per capita income growth rates are negatively correlated in a cross section does not invalidate the model. Notice that this fact makes it difficult — though not impossible, as we will see momentarily — to test the model with cross-section evidence. Ideally, one needs a cross-section of economies that cannot share ideas.

One useful laboratory for testing the model is the very long-run historical experience of the world economy, and this evidence has been explored in detail by Kremer (1993). Kremer collates population data going back one million years and documents that population growth was extremely slow prior to the Industrial Revolution, at which point it increased dramatically, as did the rate of growth of per capita income. These basic facts are con-

[17] A similar argument applies for a capital tax. However, more complicated policies are required when $\epsilon \neq 1$.

sistent with the model. A second and intriguing piece of evidence from Kremer is the natural experiment related to the last ice age. Five regions of the world — Flinders Island, Tasmania, Australia, the Americas, and the Eurasian continent — were completely isolated from one another at the end of the most recent ice age, about 10,000 years ago. These regions varied considerably in area and population from the tiny Flinders Island to the large, populous Eurasian continent. Still, 10,000 years ago, all populations were essentially hunters and gatherers, presumably at near-subsistence income levels. In contrast, by the year 1500, when large ocean-worthy ships made possible the integration of the world economy, large differences in technology levels existed among these regions. Kremer shows that the initial populations of these regions 10,000 years earlier (rank) correlate exactly with their technology levels in the year 1500. This is direct evidence of a scale effect in levels.

A final issue worth considering is the plausibility of the fertility model developed here. While it may be possible to generate something like a demographic transition using transistion dynamics, a clear prediction of the model is positive steady state population growth, at least for the parameter values considered here.[18] In contrast, demographic projections by the U.S. Bureau of the Census and the World Bank suggest that world population may stablize at some point far into the future like the 23rd century (Doyle 1997).

While the model does not necessariliy predict this behavior (and given the history of demographic projections it is not obvious how seriously one should take this prediction), the implications of a stable population are easily explored. In particular, the analysis of the production function for new ideas that leads up to equation (15) implies that the growth rates of the stock of knowledge and per capita income would asymptotically go to

[18]It would be interesting to incorporate the more detailed modeling of fertility and the demographic transition by Galor and Weil (1996) and others into this setup.

zero with a stable population. Notice that the creation of new ideas would not cease — in the simplest case with $\lambda = \phi = 0$, a constant number of researchers would create a constant number of new ideas. However, this constant number of new ideas as a fraction of the total stock of knowledge would gradually go to zero. Some readers may find this prediction to be unreasonable, but it seems difficult to judge its accuracy without actually running the experiment.[19]

5 Conclusion

Why do modern economies like the United States and Japan exhibit sustained growth in per capita income? This is one of the fundamental questions underlying research in the endogenous growth literature. At a technical level, this question can be reduced to What is the best way to characterize the linear differential equation that is a necessary element of any model that exhibits long-run growth?

To date, the endogenous growth literature has focused on two different locations for the linearity: the accumulation equations for traditional inputs like physical and human capital, and the equation governing technological progress. Neither location has proved satisfactory, both for empirical reasons and more fundamentally because at some level the linearity is simply assumed *ad hoc* rather than motivated from first principles.

This paper proposes a new source for the linearity responsible for endogenous growth, the fertility equation. A representative agent chooses to have a certain number of children, $\tilde{n}$. With N such agents in the economy, the net increase in population is given by $\dot{N} = nN$, where $n = \tilde{n} - 1$. In other words, by picking the number of children to have, individuals choose

[19] Such readers may then be interested in adding endogenous fertility to Young's (1995) model. While growth in his model would fall as fertility declined, it would asymptote to a positive rate rather than to zero.

the *proportional* rate of increase in the population. The linearity of the law of motion for population results from the biological fact of nature that people reproduce in proportion to their number. By itself, however, this linearity is not sufficient to generate per capita growth.

The second key ingredient of the model is increasing returns to scale. Following the reasoning of Romer (1990) and others, increasing returns also seems to be a fact of nature. Ideas are a central feature of the world we live in. Ideas are nonrivalrous. Nonrivalry implies increasing returns to scale. This line of reasoning, rather than placing the key linearity in the equation of motion for technological progress, is the fundamental insight of the idea-based growth models, according to the view in this paper. Endogenous fertility and increasing returns, both motivated from first principles, are the key ingredients in an explanation of sustained and endogenous per capita growth.

It is common to ask of endogenous growth models if policies can affect the long-run growth rate. The view of the mechanism underlying endogenous growth presented in this paper puts sharp limits on the channel through which such changes can occur. In the model, policies such as subsidies or taxes on research, capital accumulation, and fertility can affect the long-run growth rate. However, these effects operate through fertility: only by altering the long-run rate of population growth can standard policies affect long-run per capita growth. Because of this restriction, many policies have long-run growth effects that are counter to conventional wisdom. For example, a research subsidy will reduce long-run growth in the model, even if it is welfare-improving. In this sense, the model emphasizes that the appropriate focus of policy should be on welfare and not on long-run growth.

References

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, March 1992, *60*, 323–351.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *Review of Economic Studies*, June 1962, *29*, 153–173.

Barro, Robert J., "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 1974, *82*, 1095–1117.

___ **and Gary S. Becker**, "Fertility Choice in a Model of Economic Growth," *Econometrica*, March 1989, *57* (2), 481–501.

Becker, Gary S. and Robert J. Barro, "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics*, February 1988, *108* (1), 1–25.

Dasgupta, Partha S., "On the Concept of Optimum Population," *Review of Economic Studies*, July 1969, *36*, 295–318.

Doyle, Rodger, "By the Numbers, Global Fertility and Population," *Scientific American*, March 1997, *276*, 26.

Galor, Oded and David N. Weil, "The Gender Gap, Fertility, and Growth," *American Economic Review*, June 1996, *86* (3), 374–387.

Grossman, Gene M. and Elhanan Helpman, *Innovation and Growth in the Global Economy*, Cambridge, MA: MIT Press, 1991.

Jones, Charles I., "R&D-Based Models of Economic Growth," *Journal of Political Economy*, August 1995, *103*, 759–784.

___ , "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, May 1995, *110*, 495–525.

Judd, Kenneth L., "On the Performance of Patents," *Econometrica*, May 1985, *53*, 567–585.

Kremer, Michael, "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics*, August 1993, *108*, 681–716.

Lucas, Robert E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22* (1), 3–42.

___ , "Making a Miracle," *Econometrica*, 1993, *61*, 251–272.

Mankiw, N. Gregory, David Romer, and David Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, May 1992, *107* (2), 407–438.

Nordhaus, William D., "An Economic Theory of Technological Change," *American Economic Association Papers and Proceedings*, May 1969, *59*, 18–28.

Phelps, Edmund S., "Models of Technical Progress and the Golden Rule of Research," *Review of Economic Studies*, April 1966, *33*, 133–45.

Pitchford, John, "Population and Optimal Growth," *Econometrica*, January 1972, *40* (1), 109–136.

Raut, L.K. and T.N. Srinivasan, "Dynamics of Endogenous Growth," *Economic Theory*, 1994, *4*, 777–790.

Razin, Assaf and Uri Ben-Zion, "An Intergenerational Model of Population Growth," *American Economic Review*, December 1975, *65* (5), 923–933.

Rebelo, Sergio, "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy*, June 1991, *99*, 500–521.

Romer, Paul M., "Crazy Explanations for the Productivity Slowdown," in Stanley Fischer, ed., *NBER Macroeconomics Annual 1987*, Cambridge, MA: MIT Press, 1987.

—— , "Endogenous Technological Change," *Journal of Political Economy*, October 1990, *98* (5), S71–S102.

—— , "Comment on a paper by T.N. Srinivasan," in "Growth Theories in Light of the East Asian Experience," Chicago: University of Chicago Press, 1995.

Shell, Karl, "A Model of Inventive Activity and Capital Accumulation."

—— **, ed.**, *Essays on the Theory of Economic Growth*, Cambridge, MA: MIT Press, 1967.

Sheshinski, Eytan, "Optimal Accumulation with Learning by Doing." In Shell, ed (1967).

Solow, Robert M., "Perspectives on Growth Theory," *Journal of Economic Perspectives*, Winter 1994, *8* (1), 45–54.

Young, Alwyn, "Growth without Scale Effects," 1995. Boston University mimeo.