Innovation, Diffusion and the Distribution of Income in a Malthusian Economy

Staley, Mark
University of Ontario Institute of Technology

15. May 2008

Innovation, Diffusion and the Distribution of Income in a Malthusian Economy

Mark Staley*

May 2008

ABSTRACT

Between 5000 BCE and 1800, the population of the world grew 120-fold despite constraints on the total amount of land available for production. This paper develops a model linking population growth to increasing productivity driven by random innovation and diffusion. People are endowed with a set of skills obtained from their parents or neighbours, but those skills are imperfectly applied during their lifetimes. The resulting variation in productivity leads to a distribution of income and to a process of diffusion whereby high-income activities spread at the expense of low-income activities. An analytic formula is derived for the steady-state distribution of income. The model predicts that the rate of growth of population approaches an asymptotic limit, whereupon there are no scale effects. The model also predicts that if the rate of diffusion of knowledge is increased, the growth rate will increase.

* University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, Ontario, Canada, L1H 7K4. Email: staleymark@hotmail.com

1. Introduction

For thousands of years prior to 1800, average per-capita income was very stable and very low. Clark (2007) provides evidence to show that the standard of living of an English peasant in the year 1800 was similar to that of a hunter-gatherer living 100,000 years ago, at last when measured in terms of nutrition and longevity. According to Thomas Malthus, incomes were stagnant for so long because *"the constant effort towards population, which is found even in the most vicious societies, increases the number of people before the means of subsistence are increased"* (Malthus, 1826). The two main assumptions of Malthus' model were that the rate of population growth was increasing in per-capita income, and that there were diminishing returns to labour because land was in fixed supply. He showed that under these two assumptions income would be mean reverting. Population would also be mean reverting unless there were improvements to skills or technology that allowed more people to subsist off the same amount of land. Hence according to Malthusian reasoning, population growth in the pre-industrial era must have been driven by innovation.

Historical evidence supports a link between population growth and innovation. Phillip Hoffman has constructed an index of total factor productivity (TFP) for agricultural land in the Paris Basin between 1500 and 1800, showing a steady increase in TFP accompanied by a similar increase in the labour force over that period of time (Hoffman, 1996, Table 4.10). Clark has found a similar pattern for England between 1600 and 1800 (Clark, 2007, Figure 2.6). A recent paper by Ashraf and Galor (2008) has shown more generally that societies characterized by higher land productivity and an earlier onset of agriculture had a higher population density in the time period 1-1500 CE. Figure 1 shows world population at around 5 million in the year 5000 BCE (when agriculture was taking hold), increasing to 600 million on the eve of the industrial revolution (Kremer [1993]). In the context of a Malthusian economy this 120-fold increase in population represents an enormous amount of innovation. However in any given decade the rate of improvement would have seemed glacial. One important characteristic of Figure 1 is that the rate of population growth appears to have been independent of the *level* of population, i.e. there were no scale effects between 5000 BCE and 1800.

A simple macroeconomic model can be used to describe the phenomenon of a steadily growing population and a constant level of per capita income. Let Y, X, L stand for output, fixed land, and labour respectively. Let A stand for labour efficiency (or human capital per person). The model consists of three equations:

Cobb-Douglas Production: $Y = X^{\alpha}\left(AL\right)^{1-\alpha}$, $0 < \alpha < 1$,

Exogenous Innovation: $A = e^{gt}$,

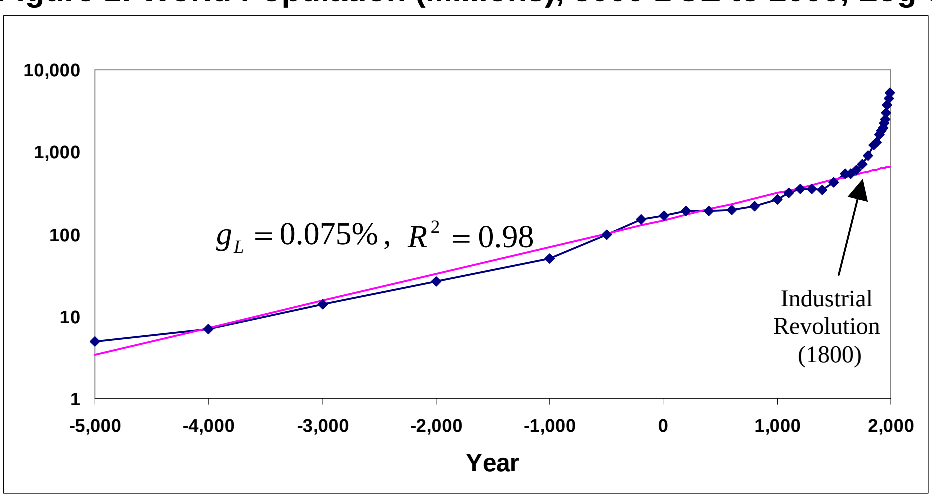
Malthusian Dynamics: $\dot{L} = \beta Y - \delta L$

The symbol β in the last equation represents the number of new labourers that survive to adulthood per unit of economic output, and δ represents the natural death rate of labourers. This last equation is analogous to the savings equation used by Solow in this 1956 model of industrial growth, but with labour substituted for capital (Solow, 1956). Following the technique used by Solow, a steady-state solution to these equations can be obtained:

Population Growth: $g_L \equiv \frac{\dot{L}}{L} = \frac{1-\alpha}{\alpha} g$,

Income Per Capita: $y \equiv \frac{Y}{L} = \frac{\delta + g_L}{\beta}$

Figure 1: World Population (Millions), 5000 BCE to 2000, Log Scale



Source: Kremer [1993]
Regression is based on data between 5000 BCE and 1800.

Given the success of the above model in capturing the essence of the pre-industrial economy one might simply stop at this point. But the model at it stands seems incomplete because it treats innovation as if it were some macroeconomic effect by which improvements in efficiency descend upon the entire population in a coordinated fashion. Intuition would suggest that innovation is more likely a local phenomenon, and improved techniques displace older techniques through a process of diffusion. There are two challenges in building a model of localized innovation and diffusion. First, one must describe how individuals come up with innovations. Second, one must avoid scale effects. Kremer (1993) presents a proto-typical model of population growth that shows how scale effects arise naturally when innovation is assumed local. In Kremer's model, each person's chance of inventing something is independent of population, so the aggregate rate of invention is proportional to population. The implication of this reasonable assumption is that the rate of population growth should be *increasing* over time. Although Kremer's data supports a pattern of accelerating population growth after 1800 (during the industrial revolution), that same set of data shows no apparent scale effects prior to 1800.

The goal of this paper is to present a *scale invariant* model of pre-industrial growth with local random innovation and diffusion. People are endowed with a set of skills obtained from their parents or neighbours, but those skills are imperfectly applied in their own lifetime. There is *no* attempt by people to purposefully improve their skills. Instead, random (directionless) variation leads to a distribution of income and to a process of diffusion whereby high-income activities spread at the expense of low-income activities. It turns out that a finite rate of diffusion puts a kind of "speed limit" on the aggregate rate of innovation and hence eliminates scale effects. As the population grows, more people discover new skills that have already been discovered elsewhere but have not yet diffused across society, i.e. they end up "re-inventing the wheel". The economy eventually settles into a steady state in which the distribution of income is stable and the rate of growth of population is independent of the level of population. A central prediction of the model is that the faster the rate of diffusion of knowledge, the faster the growth rate of population.

A key assumption of the model is that knowledge diffuses through society at a finite rate. Modern evidence shows that the diffusion of superior technologies is indeed not instantaneous, even when the benefits are seemingly clear and there are no legal barriers to adoption. One of the best-known studies of diffusion concerns the adoption of hybrid seed corn by farmers in Iowa between 1930 and 1950 (Ryan & Gross [1943]). During those decades, hybrid seed led to yields that were 20% higher than those common at the time. But farmers were conservative, tending not to switch to the new seed until they had witnessed their neighbours enjoying success. As Griliches [1957] observed, the pattern of adoption was S-shaped: there was an initial period of slow adoption, followed by a period when the rate of adoption was high, followed by a levelling-out process as the pool of potential new users shrank. Subsequent work has shown that this S-shaped pattern of diffusion is practically ubiquitous (Rogers [1995]).

The process whereby knowledge is spread through society by way of direct encounters between people as described above might be called horizontal diffusion. An alternative type of diffusion, perhaps even more important in the pre-industrial era, was that between parent and child, i.e. vertical diffusion. In a society with no public education and limited opportunities for travel, people would have learned most of their skills from their parents. Such vertical transfers of knowledge would lead to the spread of superior techniques under Malthusian conditions because members of the most productive families would leave the most offspring. And given that land was in fixed supply one might expect to see a process of selection, similar to Darwinian selection, acting to favour the people with the highest levels of skills and knowledge.[1]

The Malthusian assumptions underpinning the present model are reviewed by Galor (2005), and Galor & Weil (2000). Several recent papers have presented models of pre-industrial growth based on Malthusian assumptions (Kremer, 1993; Jones, 1999; Galor & Moav, 2002; Hansen & Prescott, 2002; Lucas, 2002). These papers are mainly concerned with the transition from Malthusian income stagnation to modern growth, while the present paper is concerned only with the Malthusian era. The role of selection in the diffusion of innovation was previously discussed by Galor & Moav, by Clark & Hamilton (2006), and by Clark (2007). Whereas those authors explored the possibility that *genetic* selection may have driven increases in productivity, the present paper considers only what may be termed *cultural* selection, i.e. changes in knowledge and

[1] From Darwin [1883]: *"...I saw, on reading Malthus on Population, that natural selection was the inevitable result of the rapid increase of all organic beings..."*. The type of selection considered here has been variously termed cultural selection or behavioural selection, to distinguish it from genetic selection (Jablonka & Lamb, 2005).

skills. Many of the results in this paper have been obtained using the tools of continuous-time stochastic calculus, originally applied to the study of economic growth by Bourguignon (1974) and Merton (1975).

The paper is organized as follows. Section 2 presents a model of population growth with random innovation and diffusion of knowledge (both horizontal and vertical). Section 3 presents an analytic formula for the distribution of income and shows how one may compute the population growth rate as a function of demographic and economic factors. Finally, section 4 summarizes the findings of this paper and suggests some possible extensions.

2. The Model

2.1 Production

The production function for each unit follows the Cobb-Douglas form:

(2.1) $$Y_i = X_i^{\alpha} \left(A_i L_i \right)^{1-\alpha} .$$

Here i labels a unit of production in which people have attained a certain level of knowledge A_i, and Y_i, X_i, and L_i stand for the levels of output, land and labour associated with that unit. We assume that the quality of land is homogeneous across all units. Note that there is no capital in this model (or equivalently, capital is tied to land or to labour in some fixed proportion[2]). A set of people may be a tribe, a manor in pre-industrial Europe, or just a family. We assume that the size of a unit is small in comparison with the whole population, so the economy is competitive. One may think of the quantity $\left(A_i L_i \right)$ as representing the amount of effective labour, or human capital. Total output for the economy is simply $Y = \sum_i Y_i$.

[2] The assumption that capital is tied to other factors seems reasonable for a pre-industrial economy. For example, draught animals were an important form of capital but required pasture for grazing so the potential for accumulation was limited.

Since income drives population in a Malthusian economy, our immediate goal is to derive an expression for per-capita income applicable to each unit. In order to do so we must first determine how land is distributed across the various units of production. Two assumptions are sufficient. First, we assume that the marginal product of land is the same across all units:

$$\frac{\partial Y_i}{\partial X_i} = \text{constant} .$$

Second, we assume that the total amount of land is fixed (normalized to 1 for convenience):

$$\sum_i X_i = 1 .$$

With these two assumptions one may show that land is distributed in proportion to human capital:

(2.2) $$X_i = \frac{A_i L_i}{AL} ,$$

where L is the total quantity of labour and A is the labour-weighted average productivity across all units. Per-capita income is then proportional to productivity:[3]

(2.3) $$y_i \equiv \frac{Y_i}{L_i} = \frac{A_i}{(AL)^\alpha}$$

The assumption of a constant marginal product of land can be justified in the context of a society where there is clear title to land and a competitive rental market. In that case, the marginal product of land is equal to the rent, and rent is the same for everyone because the quality of land is assumed homogeneous across all units of production.

In a society without formal land ownership the distribution of land is more likely determined by military strength. But even then one can argue that as long as people are rational, land will be distributed as described above. Consider the situation where there are two neighbouring units of production, one of which enjoys a high marginal product of land (labelled "H"), and the other of which has a low marginal product of land (labelled

[3] An interpretation of Equation (2.3) is that each unit earns its average product of labour, i.e. $y_i = \frac{Y}{AL} A_i$.

“L”). All units wish to expand their holdings of land because according to (2.1) that will allow them to expand output. It is profitable for a given unit to invade its neighbour only if the military budget required to defend the acquired piece of land is less than its marginal product. It turns out that unit “H” can economically expand its territory at the expense of unit “L” if it spends an amount on defence that is intermediate between the marginal products of the two units. In that case unit “L” will find it uneconomical to match the military spending of unit “H” and will be forced to retreat. A stalemate will be obtained when land is divided between the two communities such that their marginal products are the same.

There remains the question of how much of income goes to fuel population growth. An extreme Ricardian view might be that only the labour share of income fuels population growth because the remainder of income (rent) is squandered by a small land-owning elite on luxury goods and military adventures. But luxury goods makers and soldiers presumably have children, so some of that rental income will support the “effort towards population”. It is not the aim of this paper to develop a full theory of land ownership, so instead we will simply assume that all income generated by a unit of production goes to support the raising of children in that unit. Equation (2.3) can then be used as the basis for a Malthusian model of population dynamics.

2.2 Fertility

Following Hansen & Prescott (2002) we assume that the rate of growth of population in a production unit is a linear function of income:

$$\frac{\dot{L}_i}{L_i} = \mathrm{B} y_i - \delta \; . \qquad (2.4)$$

Here y_i is given by Equation (2.3) and B and δ are constants. Since the size of the labour force is proportional to total population, Equation (2.4) also describes the rate of growth of labour. The first term in Equation (2.4) then represents the rate of entry into the labour force, which is roughly equal to the number of children that survive to adulthood. Here adulthood means the ability to both work and reproduce. A natural interpretation of the first term in Equation (2.4) is that wealthy parents produce more children than poor parents.[4] The second term in Equation (2.4) represents the natural death rate of labourers.

Equation (2.4) can be derived by assuming a constant elasticity of parent's marginal utility with respect to both net consumption c (after child-rearing expenses) and the number of children n. E.g.

$$U(c,n) = \frac{\left(cn^{\phi}\right)^{1-\theta} - 1}{1-\theta}, \quad c = y - kn.$$

Here k is the expenditure required to raise a single child. For a given level of family income y, utility is maximized when $n = \mathrm{B}y$, $\mathrm{B} = \dfrac{\phi}{k(1+\phi)}$.

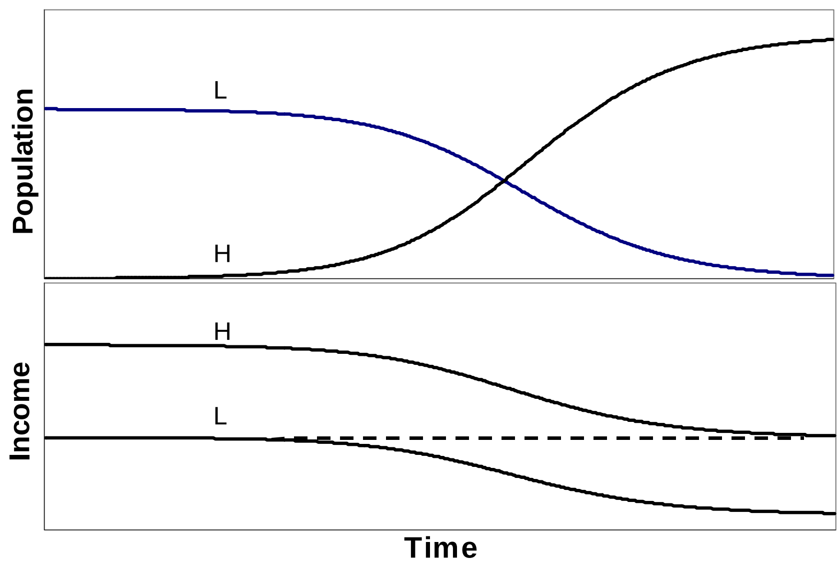
If we assume that knowledge is passed down through the generations, then Equations (2.3) and (2.4) together define a system that exhibits the characteristics of Darwinian selection. Consider a hypothetical situation where there are two types of labourers, one representing the majority (labelled "L"), and the other representing a small minority having above-average skills (labelled "H"). Figure 2 shows schematically what happens to these two populations over time.

Since the population of type "H" individuals is small at first, the equilibrium of the economy is initially dictated by the properties of type "L" individuals. The net income of type "L" individuals is just high enough to allow the population to remain stable (each couple produces on average two children that survive to reproduce). But the type "H" individuals enjoy a higher income and so are able to grow in number. The marginal product of land in type "H" communities is temporarily higher than that of the land controlled by the type "L" communities. Hence by the logic of the previous section the territory controlled by type "H" individuals grows at the expense of the territory controlled by type "L" communities until the marginal products of land are equalized. In the new equilibrium the net income of type "L" people is lower than before, so their numbers start

[4] In a paper entitled "Survival of the Richest", Clark and Hamilton provide evidence based on parish records from pre-industrial England showing a positive correlation between the number of heirs listed in wills and the total assets of testators, the later presumably a good proxy for income (Clark & Hamilton, 2006).

to shrink (e.g. the rate of infant mortality goes up). In the meantime, the population labelled "H" continues to rise, which according to Equation (2.2) triggers further expansion of territory. The process continues until type "H" individuals have taken over the economy. As expected, the higher level of population absorbs the higher income of the more productive people, to the extent that disposable per-capita income once again reverts to its original subsistence level.

Figure 2: Selection in a Malthusian Economy



The mechanism of selection just described is similar to the evolutionary mechanisms presented by Nelson & Winter (1982) but with the roles of capital and labour reversed. In Nelson & Winter's models, firms grow by reinvesting capital while competing for finite supplies of other resources such as labour. To find something even closer to the present model one must look to the theoretical population ecology literature.[5] Ecological models typically capture competitive dynamics by assuming the existence of a common limiting resource, and a population growth equation similar to (2.4), but with a "crowding term" $C(L)$, e.g.

$$\frac{\dot{L}_i}{L_i} = \beta_i C(L) - \delta \,.$$

[5] See Vandermeer & Goldberg (2003) Ch. 1, for example.

$C(L)$ is a decreasing function of the total population L, and β_i is a population specific birth parameter. We can relate Equation (2.4) to the above ecological model by mapping $A_i \to \beta_i$ and $\mathrm{B}/(AL)^{\alpha} \to C(L)$. In the model of population growth given by Equations (2.3) and (2.4), people are competing for a fixed quantity of land and the most productive people can survive crowded conditions that are too onerous for other types of people. The main difference between the ecology models and our model of Malthusian dynamics is that in the former case fertility rates are determined by genetics, whereas in our case fertility rates are determined by skills, which are nevertheless passed to descendents as if they were something like genes.

2.3 Horizontal Diffusion

Consider two populations, labelled i and j, and assume that the productivity of any person in population j is greater than that of a person in population i. That is, $A_j > A_i$. According to Equation (2.3) this productivity difference manifests itself as a difference in income, i.e. $y_j > y_i$. One would expect that if a person in population i came into contact with a person in population j and was able to observe the superior techniques used by the person in population j, then there would be a transfer of knowledge. This type of knowledge transfer can be captured using epidemic models, which generally assume that the rate of transfer between two populations is proportional to the product of the two populations (see Giroski, 2000 for a review). The resulting dynamics gives rise to an S-shaped pattern of diffusion as seen by Griliches in his famous paper on hybrid corn (Griliches, 1957). The Bass model of diffusion (Bass, 1969), widely used by marketers to forecast the spread of technology, is also based on this type of rule.

Another aspect of diffusion observed by Griliches, and also emphasized by Rogers (1995), is that the speed of diffusion appears to be proportional to the economic benefit that is obtained by switching to the superior technique or technology. This aspect of diffusion can be captured by assuming that the speed of diffusion is an increasing function of the difference in income.

We now postulate the following dynamics for the horizontal diffusion of knowledge between two populations, labelled i and j:

(2.5) $$\dot{L}_j = \nu L_i \frac{L_j}{L}\left(y_j - y_i\right),$$

(2.6) $$\dot{L}_i = -\dot{L}_j = \nu L_j \frac{L_i}{L}\left(y_i - y_j\right),$$

where L is the total population. The flow of knowledge is always from lower-income activities to higher-income activities. The constant ν captures the speed of diffusion. Note the symmetry between equations (2.5) and (2.6): only one of these equations is needed to specify the model. The intuition behind Equation (2.5) is that the rate of increase in the population with superior knowledge A_j is proportional to the number of potential learners (L_i), and is also proportional to the percentage of labourers that have already attained that level of knowledge (L_j / L). This second factor represents the likelihood that a potential learner will be neighbours with a potential "teacher", which captures the epidemic nature of diffusion.

In an economy with many different levels of productivity, the rate of diffusion away or towards a given level of knowledge can be obtained by summing the effects of diffusion over all relevant pairs of types. Hence to obtain the total rate of change of L_i, we can sum Equation (2.6) over j to obtain

(2.7) $$\dot{L}_i = \nu L_i (y_i - \overline{y}),$$

where $\overline{y}$ is the labour-weighted average wage across the economy.

It is useful at this point to place the above model of knowledge diffusion in the context of technology diffusion models. Giroski (2000) classifies diffusion models into four categories: epidemic models, probit models, density-dependent population models, and information cascade models. Epidemic models have already been discussed. Probit models postulate that units of production (firms in modern parlance) are heterogeneous in their ability to adopt new technologies. For example, a firm may adopt a new technology only if the profit in doing so exceeds some threshold, say π^*. Let's say the distribution of π^* across firms is $f(\pi^*)$. The proportion of firms adopting the technology is then equal to the area under $f(\pi^*)$ where π^* is less than the increase in profit obtained by switching to the new model. Density dependent population models include the model of selection described in the previous section. Information cascade models describe the phenomena whereby "herd mentality" may cause firms to adopt a certain technology even when there are other more profitable alternatives.

The model of diffusion described by Equations (2.5), (2.6) and (2.7) is a hybrid of epidemic and probit models. The factor $\nu L_i L_j / L$ in Equation (2.5) captures the epidemic dynamics, while the dependence on $y_j - y_i$ captures the heterogeneity of capabilities across units of production. One can derive the $y_j - y_i$ factor from a probit model by assuming that each unit of production adopts a new technique only if the accompanying increase in income exceeds some threshold. If that threshold is uniformly distributed across units, then the proportion of units adopting the given technique will be linear in $y_j - y_i$.

Note that Equation (2.7) is similar to the equation for population dynamics presented in the previous section (Equation (2.4)). The correspondence can be seen if we map $\nu \to \mathrm{B}$ and $\nu\overline{y} \to \delta$ ($\nu\overline{y}$ is a constant since $\overline{y}$ is constant in a Malthusian economy). Hence our model can also be viewed as a selection model. In epidemic models, selection acts on different variants of pathogens that are competing for hosts. If skills are something like pathogens, then these skills are "competing for people" and only the most communicable will survive, communicability in this case being related to differences in income. [6]

Finally, Equation (2.7) can be combined with Equation (2.4) to obtain the total rate of change of population of a given type:

(2.8) $$\frac{\dot{L}_i}{L_i} = By_i - \delta + \nu\left(y_i - \overline{y}\right)$$

This last equation combines the effects of vertical diffusion and horizontal diffusion.

[6] Dawkins concept of a meme comes closest to capturing the idea of skills competing for people (Dawkins, 1976).

2.4 Innovation

Our model of innovation is very simple:

(2.9) $$\frac{dA_i}{A_i} = \sigma dz_i .$$

Here σ is a constant and dz_i represents a draw from a standardized iid normal process: $dz_i \sim N(0, dt)$. There is no direction to innovation, and productivity is as likely to decrease as it is to increase. To simplify matters we assume that individuals innovate independently of one another, so the dz_i are uncorrelated across i. One complication that we will need to address later is that the definition of a group may change over time. For example, a group may start out as a single "tribe", but after the population expands, the tribe may split into two independent tribes each pursuing their own innovation according to Equation (2.9).

2.5 Summary of the Model

(2.3) Income:	$y_i \equiv \frac{Y_i}{L_i} = \frac{A_i}{(AL)^{\alpha}}, \quad A = \sum_j \frac{L_j}{L} A_j, \quad L = \sum_i L_i ,$
(2.8) Diffusion/Selection:	$\frac{\dot{L}_i}{L_i} = \mathrm{B} y_i - \delta + \nu\left(y_i - \overline{y}\right), \quad \overline{y} \equiv \sum_i \frac{L_i}{L} y_i ,$
(2.9) Innovation:	$\frac{dA_i}{A_i} = \sigma dz_i .$

3. Aggregate Growth

3.1 Simulation of the Model

Before developing an analytic model of aggregate growth we present the results of a simulation exercise, designed to replicate the pattern seen in Figure 1. The purpose of the simulation is to highlight some key properties that will need to be captured in the analytical solution. Some of the parameters of the simulation have been chosen to be consistent with historical data; others have been chosen based on plausibility. Together, they are designed to produce a rate of growth of 0.075% per year, consistent with the slope of the regression line shown in Figure 1.

First, we assume that there are a large number of “tribes”, each of which is restricted in size. Initially there are 1,000 tribes, each containing 5,000 people; hence there are 5 million people to start. Whenever a tribe grows beyond 10,000 people, it splits into two. This step is necessary to prevent any single tribe from taking over the entire economy, contradicting the assumption that there are a large number of units of production. We are assuming that as population expands in the pre-industrial world, the number of communities expands with it, instead of each community becoming larger.

Parameters: Using data from Hansen & Prescott (2002), along with our estimate of growth $g_L = 0.075\%$, we can infer that δ must be approximately $1/18$ in units of years^{-1}, which implies that in the pre-industrial era the average person could expect to live an additional 18 years upon reaching adulthood.[7] Summing (2.8) over i we have

(3.1) $$g_L \equiv \frac{\dot{L}}{L} = \mathrm{B}\bar{y} - \delta ,$$

and hence $\mathrm{B} = (g_L + \delta)/\bar{y} = 0.05625/\bar{y}$. According to figures contained in Maddison (2007), world GDP per capita prior to 1800 was roughly US \$500 in 1990 terms. Therefore $\mathrm{B} \simeq 0.05625/500 = 0.0001125$. For simplicity let us assume initially that there is no horizontal diffusion, so $\nu = 0$ and all diffusion occurs vertically through a process of

[7] Annualising Equation (15) in Hansen and Prescott, $\delta = 2(1/35 - g)$ where $g = 0.075\%$, hence $\delta = 0.0555$.

selection. The share of land in production, α, is set to 0.3. The one parameter left to be determined is σ. By a process of trial and error it was determined that $\sigma = 0.0022$ results in a rate of growth that is close to 0.075%. This value of σ implies that productivity fluctuates with a standard deviation of 0.22% per annum, or just less than 1% over the average working life of an individual.

The main finding of the simulation exercise is that the distribution of income quickly adopts the form shown in Figure 3, even when starting from an arbitrary shape. The mean of the distribution ($500) is quite stable, which translates into a constant population growth rate and hence no scale effects. Note that the mean income of $500 is close to the subsistence level of $\delta/\mathrm{B} = \$493$ (substitute $g_L = 0$ into Equation 3.1). The standard deviation of the distribution is approximately $38. Over 99% of the population has an income that lies somewhere between $400 and $600. The solid line in Figure 3 represents a normal distribution with the same mean and standard deviation as the simulated distribution. It turns out that the normal distribution fits the simulated results very well, although as we shall see in the next section the exact distribution is actually related to a Bessel function.

Although the average level of income is almost constant after a few hundred years of simulation, there is still a small residual dependence on population. Figure 4 shows that the dependence appears to be approximately of the form $a - b/N$, where a and b are constants and N is the number of tribes. As $N \to \infty$ the average income (and hence the growth rate) becomes independent of population.

Figure 3: Distribution of Income

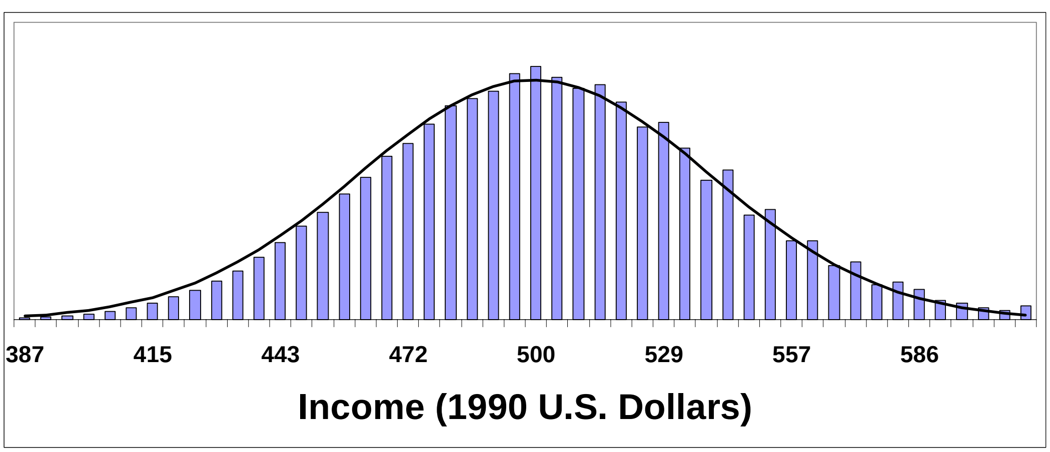
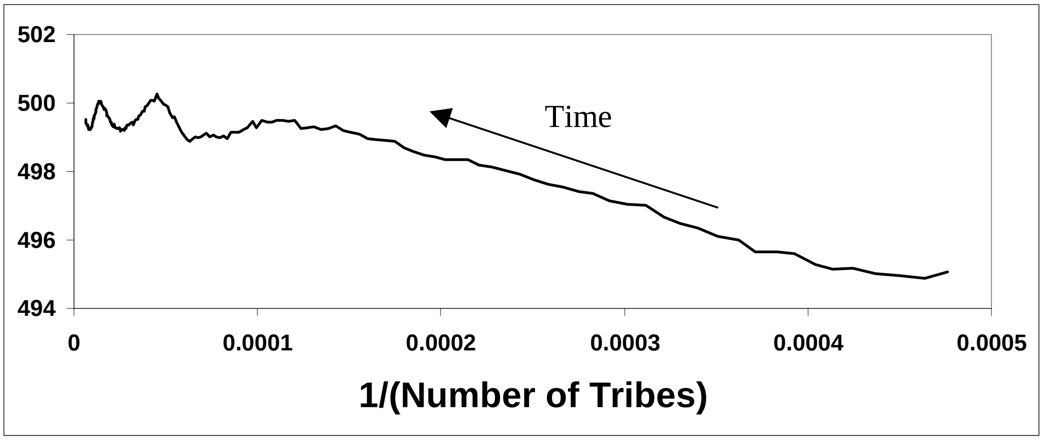


Figure 4: Dependence of Mean Income on Population



The observed dependence of the growth rate (and mean income) on the parameters B, δ, ν and σ is as follows:

1. When $\nu = 0$, the population growth rate is independent of B. However, average income is negatively correlated with B as expected in a Malthusian economy. When $\nu > 0$, the population growth rate is negatively correlated with B.

2. If one increases the death rate δ, the level of population drops and the average level of income increases, as expected in a Malthusian economy. But interestingly, the population growth rate goes up.

3. The population growth rate and the average income are both positively correlated with ν.

4. The population growth rate and the average income are both positively correlated with σ.

The last observation conforms to the expectation that in a "Darwinian" economy the rate of increase in mean productivity should be a positive function of the variance of productivity. The more variance there is, the more that selection has to operate upon. Some of the other results are surprising. For example, the second bullet point says that if the death rate increases, the rate of growth of population increases, and it is not at all obvious why this relationship should hold.

By studying the workings of the simulation one can obtain a picture of how diffusion and selection work in this economy and so gain some intuition around the relationships reported above. The picture that emerges is as follows. There are a large number of units, some of which are operating close to the frontier of knowledge, while others are further behind. The more units there are operating near the frontier, the more likely it is that one of them will accidentally discover something that increases the overall productivity of the economy. Furthermore, the faster the speed of diffusion, the more units will be located near the frontier of knowledge. Therefore an increase in the speed of diffusion should increase the growth rate. Regarding the second point, an increase in the death rate leads to a higher rate of selection (i.e. the slope of the curves in Figure 2 are steeper), and therefore a higher rate of (vertical) diffusion.

Finally, we can gain some intuition around why there are no scale effects in this model. One might expect that as the population increases, the rate of discovery should go up because there are more units drawing independent samples from the productivity distribution (dz_i). However, not everyone is operating near the frontier of knowledge. Those that are lagging the frontier are making discoveries just as fast as those that are ahead, but the laggards are effectively "re-inventing the wheel". As the economy expands this phenomenon becomes more and more common, counteracting the increased rate of discovery.

3.2 Analytic Solution

We now wish to find an expression for the growth rate of population g_L as a function of the parameters α, B, δ, ν and σ. According to Equation (3.1), the overall growth rate of population is a simple linear function of the average income $\bar{y}$, so we can direct our efforts towards finding $\bar{y}$.

In the summary of the model shown in section 2.5, we listed some differential equations for L_i and A_i; and we expressed y_i as a function of A_i. So it should be possible to find a differential equation for $\bar{y}$ and find its fixed-point solution. To carry out this procedure we make use of *Ito's lemma* (Ito [1951]):[8]

Ito's Lemma:

Given the process $dx_i = a(x_i)dt + b(x_i)dz_i$, and function $f(\{x_i\},t)$, then

$$df = \frac{\partial f}{\partial t}dt + \sum_i \frac{\partial f}{\partial x_i}\{a(x_i)dt + b(x_i)dz_i\} + \frac{1}{2}\sum_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}b(x_i)b(x_j)\rho_{ij}dt,$$

where ρ_{ij} is the correlation between dz_i and dz_j.

Since innovations are assumed to be independent, ρ_{ij} is a matrix with ones down the diagonal and zeros everywhere else, i.e. $\rho_{ij} = \delta_{ij}$. Hence

$$df = \frac{\partial f}{\partial t}dt + \sum_i \frac{\partial f}{\partial x_i}\{a(x_i)dt + b(x_i)dz_i\} + \frac{1}{2}\sum_i \frac{\partial^2 f}{\partial x_i^{\,2}}b(x_i)^2 dt$$

From Equation (2.3) we have

(3.2) $$\bar{y} = \frac{A^{1-\alpha}}{L^{\alpha}}.$$

Applying Ito's lemma to (3.2) using (2.3), (2.8) and (2.9) (see the box, Section 2.5) we can obtain $d\bar{y}$ in several steps. First:

(3.3) $$dA = A\left\{(\mathrm{B}+\nu)\frac{\sigma_y^2}{\bar{y}}dt + \sigma\sqrt{H(1/N)}dZ\right\}$$

where

$\sigma_y^2 = \frac{1}{N}\sum_i (y_i - \bar{y})^2$ is the variance of income,

$H(1/N) = \frac{\sum_i L_i^2 y_i^2}{\left(\sum_i L_i y_i\right)^2}$ is a Herfindahl index (a measure of concentration),

[8] A non-rigorous derivation of Ito's lemma can be found in Hull [2003].

$dZ \sim N(0,1)$ is a standardized normal,

and N is the number of production units (e.g. tribes, manors). One can readily see that the Herfindahl index is of order $1/N$ by scaling the number of production units by a positive factor. From Ito's lemma we then have:

(3.4) $$dA^{1-\alpha} = \frac{(1-\alpha)}{A^{\alpha}} dA - \frac{1}{2}\alpha(1-\alpha)\sigma^2 A^{1-\alpha} H\left(1/N\right) dt .$$

Combining (3.3) and (3.4) with $dL = L(\mathrm{B}\overline{y} - \delta)dt$ we finally obtain:

(3.5) $$d\overline{y} = \left\{ (1-\alpha)(\mathrm{B}+\nu)\sigma_y^2 + \alpha\overline{y}(\delta - \mathrm{B}\overline{y}) - \alpha(1-\alpha)\frac{\sigma^2}{2}\overline{y}H\left(1/N\right) \right\} dt$$
$$+(1-\alpha)\sigma\overline{y}\sqrt{H\left(1/N\right)}dZ ,$$

The equation for $\overline{y}$ is mean reverting, with the point of attraction being a negative linear function of $H(1/N)$, which explains the pattern shown in Figure 4. In the limit that $N \to \infty$, *H* vanishes, as does the stochastic term. The vanishing of the stochastic term is analogous to the elimination of unsystematic risk in a large diversified portfolio of assets. We are then left with a deterministic portion only, which simplifies to:

(3.6) $$\frac{d\overline{y}}{dt} = (1-\alpha)(\mathrm{B}+\nu)\sigma_y^2 + \alpha\overline{y}(\delta - \mathrm{B}\overline{y}) .$$

In deriving Equation (3.6) we have not considered the issue of changing group structure, such as occurred in the simulation exercise where we continuously split old tribes into new tribes. However, it turns out that this complication is irrelevant for Equation (3.6) because at any given time, quantities such as $\overline{y}$ and σ_y are invariant under splitting.

A stable fixed-point solution can be obtained by setting both sides of Equation (3.6) equal to zero.[9] At the fixed point, σ_y^2 is related to $\overline{y}$ as follows:

(3.7) $$\sigma_y^2 = \frac{\alpha}{1-\alpha}\frac{\overline{y}\left(\mathrm{B}\overline{y} - \delta\right)}{(\mathrm{B}+\nu)} .$$

[9] The fixed point is stable by inspection of the last term in Equation (3.6).

Using the parameter values listed in section 3.1, we obtain $\sigma_y =$\$37.80, which is close to the value of \$38 obtained in the simulation exercise.

In order to proceed further in deriving an expression for $\bar{y}$ in terms of the parameters of our model, we need to find another equation for σ_y^2. We can apply Ito's lemma to derive an equation for $d\sigma_y^2 / dt$ and try to find its fixed point, but it turns out that the fixed-point equation for σ_y^2 then has a term containing the third moment. Going further we could derive a whole set of recurrences relations for the higher moments, but that would only lead to an infinite regress. Clearly we need to determine the entire density function for income.

The most direct approach to finding the density function is to first derive it for each y_i, and then sum over i (weighting by L_i / L). We can sum the individual distributions because we are assuming no correlation between the various stochastic processes driving the changes in productivity. First, we apply Ito's lemma to Equation (2.3) using (2.8) and (2.9) to obtain, in the limit of an infinite number of tribes:

(3.8) $$dy_i = y_i \left\{ -\alpha(\mathrm{B}+\nu)\frac{\sigma_y^2}{\bar{y}} + \alpha(\delta - \mathrm{B}\bar{y}) \right\} dt + y_i \sigma dz_i \, .$$

This describes a simple process of geometric Brownian motion.

Next, to compute the density function for y_i we use the Fokker-Planck Equation, also known as the Kolmogorov Forward Equation (Cox & Miller, 1996):

Fokker-Planck-Kolmogorov Equation:

Given the process $dy_i = a(y_i)dt + b(y_i)dz_i$, the density function $\rho_i(y_i,t)$ satisfies

$$\frac{\partial \rho_i(y_i,t)}{\partial t} = -\frac{\partial}{\partial y_i}\left[a(y_i)\rho_i(y_i,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial y_i^2}\left[b(y_i)^2 \rho_i(y_i,t)\right],$$

Application of this Equation to (3.8) leads to the following partial differential equation for $\rho_i(y_i,t)$:

(3.9) $$\frac{\partial \rho_i}{\partial t} = \left\{ \sigma^2 + \alpha(\mathrm{B}+\nu)\frac{\sigma_y^2}{\bar{y}} - \alpha(\delta - \mathrm{B}\bar{y}) \right\} \rho_i$$

$$+ \left\{ 2\sigma^2 + \alpha(\mathrm{B}+\nu)\frac{\sigma_y^2}{\bar{y}} - \alpha(\delta - \mathrm{B}\bar{y}) \right\} y \frac{\partial \rho_i}{\partial y}$$

$$+ \frac{\sigma^2}{2} y^2 \frac{\partial^2 \rho_i}{\partial y^2}.$$

Here we have dropped the index i on y_i because the income scale is common across all units. Now define $f(y,t)$ as the distribution of income across all units of production:

$$f(y,t) \equiv \sum_i \frac{L_i(t)}{L(t)} \rho_i(y,t).$$

We may now derive a partial differential equation for f using (3.9), along with (2.8) and (3.1):

$$\frac{\partial f}{\partial t} = \left\{ (\mathrm{B}+\nu)(y-\bar{y}) + \sigma^2 + \alpha(\mathrm{B}+\nu)\frac{\sigma_y^2}{\bar{y}} - \alpha(\delta - \mathrm{B}\bar{y}) \right\} f$$

$$+ \left\{ 2\sigma^2 + \alpha(\mathrm{B}+\nu)\frac{\sigma_y^2}{\bar{y}} - \alpha(\delta - \mathrm{B}\bar{y}) \right\} y \frac{\partial f}{\partial y}$$

$$+ \frac{\sigma^2}{2} y^2 \frac{\partial^2 f}{\partial y^2}$$

The term $(\mathrm{B}+\nu)(y-\bar{y})$ in the above expression captures the effect of knowledge diffusion, while the rest of the expression is identical to (3.9).

A steady-state distribution of income is obtained when $\partial f / \partial t = 0$. Substituting for σ_y^2 from (3.7) we obtain the following ordinary differential equation for $f(y)$:

(3.10) $$0 = y^2 f'' + ayf' + (by + c) f$$

where

$$a = \frac{2}{\sigma^2}\left\{ 2\sigma^2 - \frac{\alpha}{1-\alpha}(\delta - \mathrm{B}\bar{y}) \right\}$$

$$b = \frac{2}{\sigma^2}\left(\mathrm{B} + \nu\right)$$

$$c = \frac{2}{\sigma^2}\left\{-(\mathrm{B}+\nu)\overline{y} + \sigma^2 - \frac{\alpha}{1-\alpha}(\delta - \mathrm{B}\overline{y})\right\}$$

Following the suggestion of Polyanin & Zaitsev (2003, p. 228), we make the substitutions

$$z = 2\sqrt{by} \text{ and } f(y) = z^{1-a}u(z).$$

Equation (3.10) then becomes

$$0 = z^2\frac{d^2u}{dz^2} + z\frac{du}{dz} + \left[z^2 - \gamma^2\right]u, \quad \gamma = \sqrt{(1-a)^2 - 4c}, \tag{3.11}$$

which is Bessel's equation. It has the solution

$$u(z) = C_1 J_\gamma(z) + C_2 Y_\gamma(z),$$

where J_γ and Y_γ are γ-order Bessel functions of the first and second kind respectively, and C_1, C_2 are arbitrary constants. Hence

$$f(y) = \left(2\sqrt{by}\right)^{1-a}\left\{C_1 J_\gamma\left(2\sqrt{by}\right) + C_2 Y_\gamma\left(2\sqrt{by}\right)\right\}.$$

In order to prevent $f(0)$ from blowing up, C_2 must be zero. To see why, expand $J_\gamma(z)$ near $z = 0$ (Abramowitz & Stegun, 1972 pg. 360):

$$J_\gamma(z) = \left(\frac{1}{2}z\right)^\gamma \sum_{k=0}^{\infty}\left\{\frac{1}{k!\Gamma(\gamma+k+1)}\left(-\frac{1}{4}z^2\right)^k\right\},$$

and
$$Y_\gamma(z) = \frac{J_\gamma(z)\cos(\gamma\pi) - J_{-\gamma}(z)}{\sin(\gamma\pi)}.$$

When $\gamma > 0$, $\lim_{z\to 0} J_\gamma(z) = 0$ and $\lim_{z\to 0} J_{-\gamma}(z) = \infty$, hence $\lim_{z\to 0} Y_\gamma(z) = -\infty$. So in order for the function $f(y)$ to be bounded at the origin, we must set $C_2 = 0$. Therefore our solution is

$$f(y) = C\left(2\sqrt{by}\right)^{1-a} J_\gamma\left(2\sqrt{by}\right), \tag{3.12}$$

where C is a normalization constant.

Now that we have the functional form for $f(y)$, the final step of our analysis to compute $\bar{y}$ as a function of the parameters of our model: α, B, δ, ν and σ. Since the coefficients of the function $f(y)$ themselves contain $\bar{y}$ (see (3.10)), we must solve for $\bar{y}$ using the consistency relation:

(3.13)
$$\frac{\int_D yf(y)\,dy}{\int_D f(y)\,dy} = \bar{y},$$

where D stands for the relevant domain of the function. It turns out that we need to restrict the domain to lie between $y = 0$ and the first non-zero root of the Bessel function, since the Bessel function is oscillatory.[10] The only way to proceed along these lines is to resort to numerical methods. Alternatively, the next section contains a useful approximation for $\bar{y}$ and g_L based on the assumption that the distribution of income is normal.

The Bessel function solution (3.12) implies that there is a maximum income attainable in the economy. Using the parameter values listed in Section 3.1, the root of the Bessel function turns out to be located near $707, which is at about the 99.99999 percentile of the normal distribution shown in Figure 3. This restriction on domain would seem to contradict Equation (2.9) because in principle it should be possible to obtain an arbitrarily large draw of a normal distribution, even if such a draw is very rare. Indeed, there is a slim chance that some unit will make a large discovery that pushes its income above the maximum, but evidently the number of such units as a percentage of the total is not stable when the number of units approaches infinity.

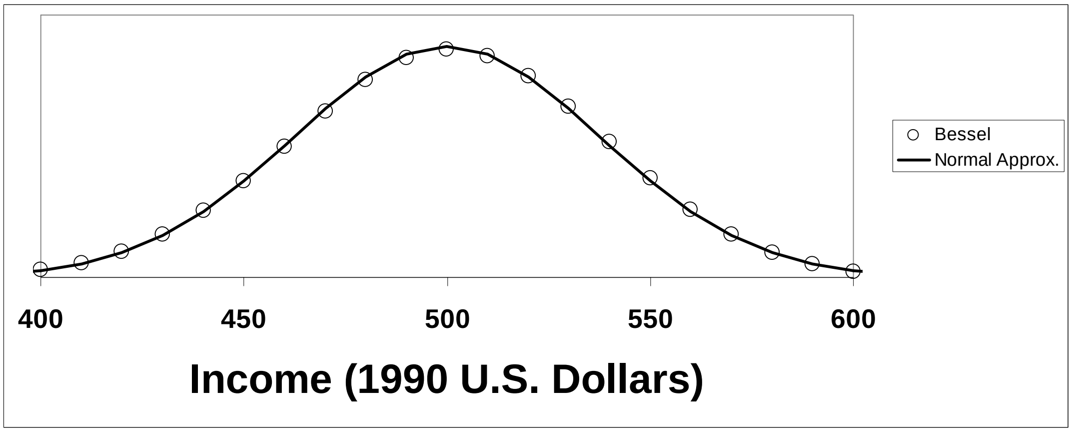
3.3 Normal approximation for $f(y)$

Recall from Figure 3 that a normal distribution fits the data from the simulation very well. Figure 5 shows the exact Bessel function solution (with Condition (3.13) verified to within

[10] The amplitudes of oscillations to the right of the first root are too small to be visible in a graph.

3 cents) overlaid on a normal distribution with the same mean of \$500 and a standard deviation of \$37.80 determined by Equation (3.7). Clearly the normal distribution is a viable base for approximation.

Figure 5: Distribution of Income: Bessel vs. Normal



It turns out that one can derive the parameters of the approximate normal distribution using Galerkin's method (Weisstein, 2008). Let us assume that $f(y)$ is approximately normal:

(3.14) $$f(y) \simeq h(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{(y-\bar{y})^2}{2\sigma_y^2}\right),$$

where σ_y is obtained from (3.7). From Equation (3.10) we have $0 = L\left(f(y)\right)$, where $L(f) = y^2 f'' + ayf' + \left(by + c\right) f$. Galerkin showed that one could approximate the solution to an ODE with some function, say $h(y)$, by solving

(3.15) $$0 = \int_{-\infty}^{\infty} h(y) L\left(h(y)\right) dy .$$

The method entails setting the weighted average error $L(h)$ across the domain of the function equal to zero, with the weight function being the target function itself. It is most often used to find the coefficients of a power series solution to a differential equation. Here we can use the method to find an approximate solution for $\bar{y}$. Substituting (3.14) into (3.15) and using

$$\int_{-\infty}^{\infty} h^2(y)dy = \frac{1}{2\sqrt{\pi}\sigma_y},$$

$$\int_{-\infty}^{\infty} (y-\bar{y})^2 h^2(y)dy = \frac{\sigma_y}{4\sqrt{\pi}},$$

$$\int_{-\infty}^{\infty} (y-\bar{y})^4 h^2(y)dy = \frac{3\sigma_y^3}{8\sqrt{\pi}}$$

$$\int_{-\infty}^{\infty} (y-\bar{y})^n h^2(y)dy = 0, \text{ where } n \text{ is odd,}$$

equation (3.15) becomes

$$0 = \frac{1}{4\sqrt{\pi}\sigma_y}\left\{\frac{\sigma^2}{4} + \frac{\alpha}{1-\alpha}\left(B\bar{y}-\delta\right) - \frac{1}{2}\frac{\sigma^2\bar{y}(B+\nu)}{\frac{\alpha}{1-\alpha}(B\bar{y}-\delta)}\right\}.$$

Substituting for σ_y using (3.7), solving for $\bar{y}$, and using $g_L = \mathrm{B}\bar{y}-\delta$ we finally obtain:

$$\text{(3.16)} \quad g_L = \frac{1-\alpha}{\alpha}\frac{\sigma^2}{4}\left[\frac{1-\alpha}{\alpha}\left(1+\frac{\nu}{B}\right)-\frac{1}{2}\right]\left\{1+\sqrt{1+\frac{8}{\sigma^2}\frac{\left(1+\frac{\nu}{B}\right)\delta}{\left[\frac{1-\alpha}{\alpha}\left(1+\frac{\nu}{B}\right)-\frac{1}{2}\right]^2}}\right\}$$

$$\simeq \frac{1-\alpha}{\alpha}\sigma\sqrt{\frac{\left(1+\nu/\mathrm{B}\right)\delta}{2}}.$$

This formula effectively captures the directional dependence of the growth rate on the parameters of our model (see list of observations in Section 3.1). A key prediction of the formula is that the growth rate is approximately linear in σ, but is approximately square-root in δ and $1+\nu/\mathrm{B}$. Table 1 shows that these dependencies are roughly born out by the simulation results (although there appears to be an upward bias to the approximated results).[11]

[11] There is a fair bit of noise in the numerical simulation results (dependent on the random number seed), which may be causing some of the discrepancy between the simulated results and the normal approximation results.

Table 1: Tests of the Numerical Approximation

Parameters				Growth Rate	
Sigma	B	Delta	v	Simulation	Normal Approx.
0.0022	0.0001125	0.0555	0	0.071%	0.086%
0.0044	0.0001125	0.0555	0	0.167%	0.171%
0.0022	0.0001125	0.08	0	0.081%	0.103%
0.0022	0.0001125	0.0555	0.0001125	0.087%	0.121%

4. Conclusion

The main premise of this paper is that it is possible to generate sustained productivity growth when individual units exhibit fluctuating productivity and there is some mechanism of diffusion that favours high-productivity units at the expense of low-productivity units. The resulting model of growth is similar to Darwin's theory of natural selection. Perhaps the strongest prediction of the model is that the distribution of income is stationary and is approximately normal, with a standard deviation that is a simple function of the average income, and of the coefficients of diffusion and demographics (Equation (3.7)).

Although the Malthusian mechanism is no longer operating in the western world, the proposed model might even have some relevance to modern industrial growth. *Horizontal* diffusion might still be acting as a selection mechanism whereby productive skills are expanding at the expense of less useful skills. Recall that one of the finding of this paper was that a finite rate of diffusion eliminates scale effects, i.e. the rate of growth of productivity is independent of the level of population. Using a model of horizontal diffusion one might be able to address the lack of observed scale effects in modern growth data (e.g. as pointed out by Jones, 1995).

In conclusion, this paper has presented a model of population growth that is consistent with the historically observed pattern between 5000 BCE and 1800. The model assumes that there are a large number of units of production that make random discoveries, which then diffuse to the rest of the population over time. Delays in the diffusion of knowledge lead to a stable distribution of income such that the resulting growth-rate of population is independent of the level of population. A related finding of the paper is that the rate of growth is an increasing function of the speed of diffusion. Finally, it has been suggested that innovation in the pre-industrial era did not flow from deliberate R&D but rather was the result of numerous random trials, from which only the most successful survived.

REFERENCES

Abramowitz, M. and Stegun, I. *Handbook of Mathematical Functions, 10th Printing*. Washington: National Bureau of Standards, U. S. Department of Commerce, 1972.

Ashraf, Q. and O. Galor. “Malthusian Population Dynamics: Theory and Evidence”, *Working Paper,* Brown University, 2008.

Bass, Frank. "A new product growth model for consumer durables". *Management Science* 15 (1969): 215-227.

Bourguignon, Francois, ”A Particular Class of Continuous-Time Stochastic Growth Models”, *Journal of Economic Theory* 9 (1974): 141-158.

Clark, Gregory and Gillian Hamilton. “Survival of the Richest. The Malthusian Mechanism in Pre-industrial England”, *Journal of Economic History*, 66 (2006):1-30.

Clark, Gregory. A *Farewell to Alms: A Brief Economic History of the World*. Princeton: Princeton University Press, 2007.

Cox, D. and H. Miller. *The Theory of Stochastic Processes*. London: Chapman and Hall, 1996.

Darwin, Charles. *The Variation of Plants and Animals under Domestication,* 2nd ed., New York: D. Appleton & Co., 1883.

Dawkins, Richard. *The Selfish Gene.* Oxford: Oxford university Press, 1976.

Galor, O. “From Stagnation to Growth: Unified Growth Theory”, in *Handbook of Economic Growth,* ed. Aghion, A & S. Durauf. Amsterdam: Elsevier B.V. 2005.

Galor, O., Moav, O. “Natural Selection and the Origin of Economic Growth”, *Quarterly Journal of Economics,* 117 (2002); 1133-1191.

Galor, O. and D. Weil, “Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond”, *The American Economic Review* 90 (2000): 806-828.

Giroski, P. “Models of Technology Diffusion”, *Research Policy* 29 (2000): 603-625.

Griliches, Zvi. "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25 (1957): 501-522

Hansen, G., Prescott, E. “Malthus to Solow”, *The American Economic Review*, 92 (2002): 1205-1217.

Hoffman, P. *Growth in a Traditional Society*. Princeton: Princeton University Press, 1996.

Hull, John. *Options, Futures & Other Derivatives*, 5th ed., New Jersey: Prentice Hall, 2003.

Ito, K. “On Stochastic Differential Equations.” *Memoirs, American Mathematical Society*, 4 (1951): 1-51.

Jablonka, E., Lamb, M. *Evolution in Four Dimensions*, Cambridge, Mass: MIT Press, 2005.

Jones, Charles. “Time Series Tests of Endogenous Growth Models", *Quarterly Journal of Economics* 110 (1995): 495-525.

______, “Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run”, Stanford University Report, 1999.

Kremer, Michael. “Population Growth and Technological Change: One Million B.C. to 1990”, *Quarterly Journal of Economics* (August 1993): 681-716.

Maddison, Angus. *The Contours of the World Economy 1-2030 AD.* Oxford: Oxford University Press, 2007. Data available at http://www.ggdc.net/Maddison/ .

Malthus, Thomas. *An Essay on the Principle of Population, 6th edition,* 1826.

Merton, Robert. “An Asymptotic Theory of Growth Under Uncertainty”, *Review of Economic Studies* 42 (1975): 375-393.

Nelson, Richard & Sidney Winter. *An Evolutionary Theory of Economic Change*, Cambridge, Mass.: Harvard University Press, 1982.

Polyanin, A. and Zaitsev, V. *Handbook of Exact Solutions for Ordinary Differential Equations, Second Ed.* Boca Raton: CRC Press, 2003.

Rogers, Everett. *Diffusion of Innovations*, 4^{th} ed. New York: Simon & Schuster Inc., 1995.

Ryan, Bryce, and Neal, C. Gross. “The Diffusion of Hybrid Seed Corn in Two Iowa Communities”, *Rural Sociology* 8, (1943):15-24.

Solow, R. “A Contribution to the Theory of Economic Growth”, *Quarterly Journal of Economics* 70, (1956): 65-94.

Vandermeer, J. and Goldberg, D. “Population Ecology”, Princeton: Princeton University Press, 2003.

Weisstein, Eric W. "Galerkin Method." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/GalerkinMethod.html, 2008.