The Transition from Stagnation to Growth: An Adaptive Learning Approach

Author(s): Jasmina Arifovic, James Bullard and John Duffy
Source: *Journal of Economic Growth*, Vol. 2, No. 2 (Jun., 1997), pp. 185-209
Published by: Springer
Stable URL: http://www.jstor.org/stable/40215955
Accessed: 13-03-2018 14:33 UTC

Journal of Economic Growth, 2: 185–209 (June 1997)

The Transition from Stagnation to Growth: An Adaptive Learning Approach

JASMINA ARIFOVIC

Simon Fraser University

JAMES BULLARD

FRB–St. Louis

JOHN DUFFY

University of Pittsburgh

This article develops the first model in which, consistent with the empirical evidence, the transition from stagnation to economic growth is a very long endogenous process. The model has one steady state with a low and stagnant level of income per capita and another steady state with a high and growing level of income per capita. Both of these steady states are locally stable under the perfect foresight assumption. We relax the perfect foresight assumption and introduce adaptive learning into this environment. Learning acts as an equilibrium selection criterion and provides an interesting transition dynamic between steady states. We find that for sufficiently low initial values of human capital—values that would tend to characterize preindustrial economies—the system under learning spends a long period of time (an epoch) in the neighborhood of the low-income steady state before finally transitioning to a neighborhood of the high-income steady state. We argue that this type of transition dynamic provides a good characterization of the economic growth and development patterns that have been observed across countries.

Keywords: growth, human capital, development, transition, learning, genetic algorithm

JEL classification: O40, D83

1. Introduction

1.1. Development Facts

Prior to industrialization, all of today's highly developed economies experienced very long periods, *epochs*, of relatively low and stagnant growth in per capita income. Maddison (1982, table 1.2) reports average annual compound growth rates in per capita GDP for sixteen of today's highly developed countries.[1] These growth rates were 0.0 percent for the years 500 to 1500, 0.1 percent for the years 1500 to 1700, and 0.2 percent for the years 1700 to 1820. It was only after industrialization, during the period 1820 to 1980, that these countries achieved a significantly higher average annual compound growth rate of 1.6

percent. While these data are highly aggregated and necessarily involve some guesswork, few economists would question the picture they paint.

Considering the more recent data, the dominant fact is that there is a large and persistent disparity in the levels of per capita income across nations (see, for example, Parente and Prescott, 1993; Durlauf and Johnson, 1995; and Quah, 1996a, 1996b). Parente and Prescott (1993), for example, use the Summers and Heston (1991) data set and report that for a sample of 102 countries over the years 1960 to 1985, per capita income in the richest 5 percent of the countries was about twenty-nine times per capita income in the poorest 5 percent of countries. The poor countries grew, on average, about as fast as the rich countries, so that this disparity has remained roughly constant over the 1960 to 1985 period. Durlauf and Johnson (1995) and Quah (1996a, 1996b) use the Summers and Heston data set to ask whether differences in cross-country growth experiences might be due to the existence of multiple steady states in per capita output. These authors show that the possibility of multiple steady states cannot be rejected.

In an effort to explain sustained differences in growth rates across economies across time, and also to explain the vast differences in levels of per capita income across nations that we observe today, a number of authors have recently expanded on the endogenous growth literature pioneered by Romer (1986) and Lucas (1988) by building models that possess multiple steady states for the growth rate of per capita output.[2] In these models, low-growth steady states, sometimes referred to as *poverty traps*, are used to characterize preindustrial or less developed economies.[3] These low-growth steady states coexist with high-growth steady states that are used to characterize industrialized or highly developed countries. While these models have certain advantages over the one-sector neoclassical growth model in the sense that they allow for sustained differences in growth rates across economies, this improvement comes at some expense: these models cannot explain how countries initially in poverty traps are ever able to make the transition to a high-development steady state.[4] Indeed, poverty traps are modeled as absorbing states from which no economy can escape. Furthermore, it is some exogenous factor, typically history or expectations, that determines whether a country will be at the low or high development steady state for all time. Yet, as Maddison's (1982) data clearly reveals, sixteen of today's most highly developed economies were in a poverty trap for many hundreds of years! These countries were nevertheless eventually able to industrialize and achieve a higher state of development.

In this article we study a model that gives rise to sustained differences in growth rates across countries for long periods of time but that also allows countries that are initially near or at low-growth steady states to eventually make the transition to high-growth steady states. The model can also account for the phenomenon that countries with similar initial conditions may experience quite different development paths, so that an observer of the world situation at a point in time might see countries with vastly different levels of per capita income.

1.2. Summary of the Model

We study the effect of adaptive learning behavior on transitional dynamics in a growth model with human capital and threshold externalities along the lines of Azariadis and

Drazen (1990). The model possesses multiple stationary equilibria, which makes it an attractive framework in which to study transition dynamics.

In this model, physical capital is accumulated in a standard way, but human capital accumulation is subject to increasing returns. Agents make two decisions when young: how much to save by renting physical capital to firms and how much to invest in training. The returns to training depend positively on the economywide average level of human capital. The model admits two steady states. The first is associated with low and stagnant levels for physical capital, human capital, and output per capita and is characterized by agents who choose not to invest in training when young. We call this the *low-income steady state*. The second steady state has higher and growing levels for these per capita variables and is characterized by agents who choose to devote a positive fraction of their available time endowment to training when young. We call this the *high-income steady state*. Each of these steady states is locally stable under a perfect foresight assumption so that, in particular, the low-income outcome is an absorbing state.[5]

Our innovation is to introduce adaptive learning behavior into this environment. We drop the assumption that agents have perfect foresight and instead assume that agents must learn which decision rules return the highest utility in the environment they face. We model learning using Holland's (1975) genetic algorithm, a stochastic, directed search algorithm based on principles of population genetics.[6] We interpret genetic algorithm learning as a useful representation of trial-and-error learning that has important advantages over many other models in the literature—chief among these for our purposes is that the genetic algorithm offers a natural model for experimentation by agents. We conduct computational experiments in order to characterize how a population of heterogenous agents might eventually find its way to the high-development steady state.

1.3. Main Findings

Our main finding is that for initially low levels of human capital per capita–levels that would tend to characterize preindustrial economies—our population of artificial agents spends many generations (an epoch) in a neighborhood of the low-income steady state before finally making the transition to the high-income steady state. We argue that this provides an account of the development fact documented by Maddison (1982) that today's richest countries were once stagnant for hundreds of years. We further demonstrate that initially identical economies might have very different development experiences in this model, in the sense that industrialization might occur at radically different times. The timing is important since different dates of industrialization imply very different postindustrialization levels of per capita income across economies in this model. We argue that this result helps explain another development fact, the present wide and persistent disparity in levels of per capita income that has been documented by Parente and Prescott (1993), Durlauf and Johnson (1995), and Quah (1996a, 1996b).

1.4. Recent Related Literature

The recent literature on growth and development is large and cannot be effectively summarized here. But the idea of multiple stationary equilibria has been a popular theme, and important contributions include Murphy, Shleifer, and Vishny (1989) on formalizing a big-push argument; Becker, Murphy, and Tamura (1990) on how fertility and human capital accumulation might interact to influence development; and Azariadis and Drazen (1990) on formalizing a threshold argument. We work in this article within the latter framework, but our approach could in principle be applied to describe transitions in these other frameworks that emphasize alternative mechanisms.

The study of transition dynamics in models with multiple steady-state equilibria has received relatively little attention. Goodfriend and McDermott (1995) build an endogenous growth model that involves transitions from premarket to market and from preindustrialized to highly developed economies. Their model has multiple stationary equilibria, but this fact plays an important role only in their explanation of the transition from premarket to market economies; their explanation of industrialization relies on a single, evolving steady state. Our approach might be useful in explaining the former transition. Galor and Tsiddon (1997) also have a model in which the economy is characterized by multiple steady-state equilibria in the short run. In their model, endogenous technological change is induced by an increase in the average level of human capital and eventually causes the structure of the dynamical system to change in a way that eliminates the poverty trap. Consequently the economy converges in the long run to a unique high-income steady-state equilibrium. While our model leads to results that are qualitatively similar to those found in Galor and Tsiddon (1997), our approach is different. In particular, the structure of the dynamical system in our model does not undergo any change that would eliminate the possibility of a poverty trap. Indeed, a poverty trap and a high-income steady state *always coexist* as possible stationary equilibria in our model.

This article is also related to the macroeconomics learning literature, which has recently been surveyed by Sargent (1993). One aim of this literature has been to use learning processes to select equilibria in models with multiple rational expectations equilibria. Our analysis is relatively novel in this literature in that our model involves capital accumulation. In addition, agents in our model are learning, simultaneously, about *two* decision rules—how much to save and how much to invest in training—in contrast to previous learning analyses, where agents are typically concerned with learning about a single decision rule.

1.5. Modeling Learning Behavior Using Genetic Algorithms

We model adaptive learning behavior using Holland's (1975) genetic algorithm—a stochastic directed search algorithm based on principles of population genetics. Genetic algorithms operate on a population of candidate solutions to some well-defined problem. Following each iteration of the algorithm, candidate solutions are evaluated for their performance and assigned a fitness value. Solutions with relatively high fitness values are more likely to remain in the next "generation" of candidate solutions than are solutions with relatively low fitness values. This process captures the notion of survival of the fittest (natural selection). The algorithm then uses the highly fit candidate solutions to breed new candidate solutions using naturally occurring genetic operations such as crossover (recombination) and mutation. A more detailed description of the genetic algorithm is provided later in Section 3.

In the context of our model, we interpret the genetic algorithm as describing a process of trial-and-error learning by a population of heterogenous, artificial economic agents. Genetic algorithms and other computational techniques involving artificial intelligence are increasingly being employed by economists as a way of modeling the behavior of economic agents. A partial list of recent references includes Andreoni and Miller (1995), Arifovic (1994, 1995, 1996), Arthur et al. (1996), Binmore and Samuelson (1992), Bullard and Duffy (1996a, 1996b), Durlauf (1995), Holland and Miller (1991), Marimon, McGrattan, and Sargent (1990), Miller (1996), Routledge (1994), Sargent (1993), Tesfatsion (1995), and Wright (1995).

We chose to model learning behavior using the genetic algorithm because it has some important advantages relative to other adaptive learning models that can be found in the literature. First, there is considerable heterogeneity across agents, a feature not often encountered in the learning literature to date.[7] Second, the information requirements on agents are minimal, as they need to know very little to function well in the economy. Third, the genetic algorithm offers a natural model for experimentation by agents with alternative decision rules, an important characteristic of learning also rarely modeled in competitive general equilibrium environments in the literature to date. Fourth, the heterogeneity of beliefs allows parallel processing to be an important feature of the economy. That is, some agents are trying one decision rule while other agents are trying other decision rules, with the better decision rules propagating and the poorer ones dying out. We think this is closely akin to what goes on in actual economies, where communication among agents encourages successful strategies to be quickly copied and unsuccessful ones to be discarded. Fifth, genetic algorithm learning has been shown in other research (see, for example, Arifovic, 1994, 1995, 1996) to successfully mimic the behavior of human subjects in controlled laboratory settings. And finally, the initial heterogeneity of the population allows us to initialize the system randomly, so that we are able to obtain some sense of the "global" properties of our system under learning as opposed to the local analysis that is often employed in the learning literature.[8] These features suggest that genetic-algorithm-based models of learning have interesting economic content.

The rest of the article is organized as follows. In Section 2 we outline the model that we employ in the rest of the article. We close the model under perfect foresight and characterize the set of stationary equilibria. In Section 3, we introduce our genetic-algorithm-based learning algorithm. Section 4 explains the design and results of our sets of computational experiments, and Section 5 concludes.

2. A Model of Growth and Development

2.1. *Preferences and Technology*

We use a version of a model of economic growth and development due to Azariadis and Drazen (1990). Time t is discrete and takes on integer values on the real line. There is a single, perishable good that is both consumed and used as an input into production. Agents in this economy live for two periods, which we label *young* and *old*. At every date t there is a total population of $2N$ agents, where N is a positive integer, with the population equally divided between young and old. There is no population growth. We use the notation that subscripts denote birthdates and parentheses denote real time, while individual agents within a generation are indexed by a superscript $i \in (1, 2, \ldots, N)$. Aggregate variables have no subscript or superscript.

Agents are endowed with one unit of time at every date t. During the first period of life, young agents may choose to spend some fraction, $\tau_t^i(t) \in [0, 1)$, of their time endowment in training. There is a common training technology, denoted $h(\tau_t^i(t), x(t))$, which all agents can access, where the variable $x(t)$ is the average quality of labor of both the young and the old at time t:

$$x(t) = \frac{1}{2}\left[\frac{1}{N}\sum_{i=1}^{N} x_t^i(t) + \frac{1}{N}\sum_{j=1}^{N} x_{t-1}^j(t)\right].$$

This variable is measured as efficiency units per worker. An individual agent can devote time to training when young in order to receive more efficiency units in the second period of life via

$$x_t^i(t+1) = h(\tau_t^i(t), x(t))x(t).$$

The key feature of the model is that the individual agent's return to training depends positively on the economywide average level of efficiency units. We follow Azariadis and Drazen (1990) and specify $h(\cdot)$ as

$$h(\tau_t^i(t), x(t)) = 1 + \gamma(x(t))\tau_t^i(t).$$

However, we depart from Azariadis and Drazen (1990) in that we use a specific parametric form for $\gamma(\cdot)$, the private yield on human capital. In particular, we use the sigmoid function

$$\gamma(x(t)) = \frac{\lambda}{1 + e^{-x(t)}} - \frac{\lambda}{2}$$

which is strictly increasing in $x(t)$ and implies the bounds given by $\gamma(0) = 0$ and

$$\lim_{x(t)\to\infty} \gamma(x(t)) = \frac{\lambda}{2} \equiv \hat{\gamma}.$$

Each young born at date t inherits the average level of efficiency units, $x(t)$, that was determined by the decisions of agents born at date $t - 1$. Young agents combine this endowment with a training decision $\tau_t^i(t)$ in order to receive $x_t^i(t + 1)$. Because we allow within generation heterogeneity in the decision variable $\tau_t^i(t)$, the accumulation equation for $x(t)$ is given by

$$x(t + 1) = x(t)[1 + \gamma(x(t))\bar{\tau}(t)],$$

where $\bar{\tau} = \frac{1}{N}\sum_{i=1}^{N} \tau_t^i(t)$.

Output per unit of effective labor is produced according to a neoclassical production function, which we specify as

$$f(k(t)) = k(t)^{\alpha},$$

where $\alpha \in (0, 1)$ and $k(t)$ is the capital-to-effective labor ratio.[9] Effective aggregate labor is given by

$$L(t) = \left[N - \sum_{i=1}^{N} \tau_t^i(t)\right] x(t) + \sum_{i=1}^{N} x_{t-1}^i(t),$$

so that

$$k(t) = \frac{K(t)}{[N - \sum_{i=1}^{N} \tau_t^i(t)]x(t) + \sum_{i=1}^{N} x_{t-1}^i(t)},$$

where $K(t)$ denotes the aggregate physical capital stock. The rental rate on physical capital and the wage are given by, respectively,

$$\begin{aligned} r(t) &= \alpha k(t)^{\alpha-1} \\ w(t) &= (1-\alpha)k(t)^{\alpha}. \end{aligned}$$

There is also a consumption loan market with gross rate of interest denoted $R(t)$. Arbitrage equates the rate of return to renting physical capital with the rate of return on consumption loans via $R(t) = r(t+1) + 1 - \delta$, where δ is the net depreciation rate on physical capital. In this article we assume $\delta = 1$.[10]

All agents in this economy have the same preferences, $U = \ln c_t^i(t) + \ln c_t^i(t+1)$. Furthermore, all agents face the same lifetime budget constraint:

$$c_t^i(t) + \frac{c_t^i(t+1)}{R(t)} \le (1 - \tau_t^i(t))x(t)w(t) + \frac{[1 + \gamma(x(t))\tau_t^i(t)]x(t)w(t+1)}{R(t)}.$$

2.2. *Equilibria Under Perfect Foresight*

In this subsection, we assume that agents have perfect foresight. Combining the first-order conditions with the budget constraint, and making use of the definitions for $w(t)$ and $R(t)$, the individual young agent's optimal savings decision can be written as

$$s_t^i(t) = \frac{(1 - \tau_t^i(t))x(t)(1-\alpha)}{2}k(t)^{\alpha} - \frac{[1 + \gamma(x(t))\tau_t^i(t)]x(t)(1-\alpha)}{2\alpha}k(t+1).$$

Young agents are equally endowed with $x(t)$, and under perfect foresight they all make the same choices for $\tau_t^i(t)$, which we call $\tau(t)$. Thus, aggregate saving is given by $S(t) = Ns_t^i(t)$. The market clearing condition is that $K(t+1) = S(t)$. Some manipulation yields

$$k(t+1) = \frac{(1 - \tau(t))\alpha(1-\alpha)}{[1 + \gamma(x(t))\tau(t)][2\alpha(2 - \tau(t+1)) + (1-\alpha)]}k(t)^{\alpha}. \tag{1}$$

We now consider steady states of this system.

First, suppose that $\tau(t) = \tau_\ell^* = 0\ \forall t$. In this case, $x(t)$ must be constant for all t. It follows from (1) that in this case

$$k_\ell^* = \left[\frac{\alpha(1-\alpha)}{1+3\alpha}\right]^{\frac{1}{1-\alpha}}.$$

The pair (τ_ℓ^*, k_ℓ^*) is the low-income steady state of our system.

Next, suppose that $\tau(t) = \tau \neq 0$. In this case, $x(t)$ is growing so that for t large enough $\gamma(t) \to \hat{\gamma}$, and furthermore arbitrage requires that $R = \hat{\gamma} = \alpha k^{\alpha-1}$. Then

$$k_h^* = \left(\frac{\alpha}{\hat{\gamma}}\right)^{\frac{1}{1-\alpha}},$$

and it follows from (1) and τ_h^* must solve

$$\left(\frac{\alpha}{\hat{\gamma}}\right)^{\frac{1}{1-\alpha}} = \left[\frac{(1-\tau)\alpha(1-\alpha)}{[1+\hat{\gamma}\tau][3\alpha - 2\alpha\tau + 1]}\right]^{\frac{1}{1-\alpha}}.$$

This is a quadratic in τ, but only one of the two roots is feasible (that is, there is only one value for $\tau \in [0, 1)$), and this is the root we choose for τ_h^*. The pair (τ_h^*, k_h^*) constitutes the high-income steady state in this system.

It is straightforward to show that the low-income steady state is locally stable in the perfect foresight dynamics and that the high-income steady state is saddlepath stable. Azariadis and Drazen (1990) argued that initial conditions would determine which steady state a nation might ultimately achieve, and that given a sufficiently diverse set of initial conditions, an observer might see nations in persistently low- as well as persistently high-growth equilibria. They argued that this prediction matches elements of the current world situation.

3. Learning

3.1. Heterogeneous Agents

We alter this model by assuming that individuals born at time t must learn about which decision rules work best in this environment. The agents are now initially heterogeneous with respect to, first, the fraction of time that they spend in training, $\tau_t^i(t) \in [0, 1)$, and second, the fraction of their time t wealth that they save. If we denote this savings fraction by $\phi_t^i(t) \in [0, 1)$, we can write a typical agent's youthful savings as

$$s_t^i(t) = \phi_t^i(t)w(t)(1 - \tau_t^i(t))x(t).$$

We model learning using a genetic algorithm, which we view as a useful model of trial-and-error learning. The genetic algorithm acts on a population of chromosomes, or strings, which are typically binary representations of important variables in the system to be studied. In our application, each binary string completely characterizes the decision rules of an individual agent. Strings are evaluated according to a fitness criterion, which in economic models is naturally taken to be a utility function. An iteration of the algorithm involves the application of genetic operators. The first operator is reproduction: strings are evaluated for fitness, and the better strings are propagated, while the poorer strings are eliminated. A second operator is crossover: new strings are created by splicing parts of existing strings together. A third operator is mutation, with which very small portions of strings are altered with small probability. Over time, the algorithm is expected to evolve strings that have, on average, higher fitness than previous generations of strings.

3.2. *Representation*

As a preparatory step to implementation of the genetic algorithm, we encode the decision rules of the entire population of $2N$ agents using binary (bit) strings. Each agent's two decision variables, $\tau_t^i(t)$ and $\phi_t^i(t)$, are encoded in a single bit string of length $\ell > 0$, where ℓ is an even integer. The first $\ell/2$ bits encode the agent's $\tau_t^i(t)$ decision and the next $\ell/2$ bits encode the agent's $\phi_t^i(t)$ decision. The initial $2N$ population of bit strings is randomly generated with each bit position in a string set equal to a zero or a one with probability .5.

To illustrate how bit strings are decoded to obtain individual's decision rules, consider an example where $\ell = 30$. An individual agent's decision string might look like this:

$$000101010011011010001101110101.$$

The first and last $30/2 = 15$ bits are decoded to obtain two base-ten integer values:

$$\underbrace{000101010011011}_{2715}\ \underbrace{010001101110101}_{9077}.$$

These integers are then divided by the maximum integer value possible, a string with 15 bits all equal to 1, plus one, which is $2^{15} = 32768$:[11]

$$\tfrac{2715}{32768} = .0828552 = \tau_t^i(t), \qquad \tfrac{9077}{32768} = .277008 = \phi_t^i(t).$$

Once we have ϕ and τ values for each of the N young agents, we can calculate each of these agent's savings decisions, $s_t^i(t)$, and we can find *aggregate savings*:

$$S(t) = \sum_{i=1}^{N} s_t^i(t).$$

From the market-clearing condition, we then find the capital-to-effective labor ratio, and we use that in turn to determine the interest rate and the wage. We can use this information to evaluate which decision rules are performing better and subsequently update the population of strings using the genetic algorithm.

3.3. *Fitness*

In the artificial intelligence literature, fitness measures how well a string performs relative to other strings. Our criterion is lifetime utility $U^i = \ln c_t^i(t) + \ln c_t^i(t+1)$. We want to be able to measure the fitness of any string in the system at time t. In order to do this, we ask the following question of each string: how well would this string have performed if it had been in use in the previous period?[12] We view the individual agent as atomistic and therefore incapable of significantly altering the level of endogenous aggregate variables in the system. Accordingly, we use past data from the system on the interest rate, the level of human capital, and the wages that the string would have faced if it had been in use in the previous period. From this we can determine how much consumption and therefore how much utility a particular string would have garnered had it been in use in the previous period. This utility level constitutes the fitness of a string.

3.4. *Genetic Operators*

3.4.1. Reproduction. At the end of period t, we have to determine the next generation of N young agents who will be born at time $t+1$, the newborns. The first step in this process is to apply the reproduction or selection operator of the genetic algorithm. The reproduction operation involves N *binary tournaments*. Each binary tournament is conducted as follows. First, choose two strings at random with replacement from the entire population of $2N$ strings—those belonging to both young and old agents—that were in use at time t. Next, compare the fitness values of the two chosen strings; the winner of the tournament is the string with the higher fitness value. A copy of this string is made and is placed in the set of strings that are candidates to become the strings used by the next generation of young agents. This binary tournament process is then repeated $N - 1$ more times yielding a set of N decision rules that are, on average, more fit than the decision rules in use at time t.

3.4.2. Mutation. Following the reproduction operation, we subject all N of the candidate strings that were winners of the N binary tournaments to some mutation. Each of the L bits in each of the N strings is independently subjected to mutation with some small, fixed probability $p^m > 0$. With probability p^m, an individual bit value is changed from $b = 0, 1$ to $1 - b$; with probability $1 - p^m$ the bit value is not changed.

3.4.3. Crossover. The final operator in the genetic algorithm is *crossover*. The crossover operator works on the population of strings that result from selection and mutation. First, each of these N strings is randomly paired with another string. For example, we might have a pairing between the following two strings:

010101000101110101110010111101

000101100101101001101100101111.

With some fixed probability, $p^c > 0$, two random integers are drawn, draw1, draw2 $\in [1, \ell/2)$. Using these numbers, the two strings are then cut at two points—one point within the first $\ell/2$ bits and one point within the last $\ell/2$ bits. For example, if draw1 $= 3$, and draw2 = 9, the two strings in our example would be cut as follows:

010|101000101110   101110010|111101

000|101100101101   001101100|101111.

The string portions to the right of each cut point would then be swapped (the substrings representing each decision are kept separate), and the two strings are then recombined:

010101100101101101110010101111

000101000101110001101100111101.

The result is two new strings, possibly representing decision rules that have never appeared in the system before.[13] The N strings resulting from selection, mutation, and crossover become the new young generation alive at time $t + 1$. The young agents alive at time t become the old agents alive at time $t + 1$, and the old agents alive at time t cease to exist (their strings are deleted). The process is repeated in order to generate a time series for the artificial economy.

3.5. *Interpreting Genetic Operators*

The reproduction, mutation, and crossover operators have a simple economic interpretation. Being *born* in this economy means leaving one's formative years and entering the productive portion of one's life. These newborn agents just leaving their formative years initially have no plans for the future: they are blank slates. They acquire the decision rules they will need by communicating with a few other members of society, those either one or two generations ahead of them. This communication is modeled via the reproduction operator. In our implementation, each newborn agent communicates with two randomly selected members of the society. The newborns evaluate the decision rules that belong to these two older agents by calculating how much utility the rules would have delivered had they been in use one period in the past. Each newborn then copies the decision rule of the two that would have delivered the most utility. This completes the first step in attaching a decision rule to each of the incoming members of the society. But the newborns communicate further when they talk with each other and contemplate alternative decision rules that might not be in use in the society at that time—that is, the newborns conduct a mental experiment with other possible decision rules. This additional communication is modeled via the crossover and mutation operators. In our implementation, the newborns are paired, and each pair creates two new decision rules by combining parts of their existing rules and also by randomly changing small parts of the decision rules. Thus the incoming generation learns from the experience of the agents older than themselves and can also be innovative in introducing new decision rules into society.

4. Design of Computational Experiments and Results

4.1. Calibration

In order to examine the behavior of our genetic-algorithm-based learning system, we conducted a large number of computational experiments. These experiments required us to choose parameterizations and initial conditions for our model, which we now describe.

There is a single parameter in the preferences and technology portion of the model that must be set: physical capital's share of output, α. We set $\alpha = .36$, a value that can be derived from postwar U.S. national income and product accounts, where consumer durables are counted as capital. By using this value, the high-income steady state of the model is consistent with postwar experience on physical capital's share in the U.S. economy.

A single parameter, λ, controls the returns to investing in human capital. These returns are partly endogenous since they depend on $x(t)$, but for large $x(t)$, $\hat{\gamma} = \frac{\lambda}{2}$. We set $\lambda = 50$, implying $\hat{\gamma} = 25$. This choice implies an endogenously determined high-income steady-state value for the fraction of time devoted to training of approximately $\tau^* + h = .22$. If we interpret the time period in the model as being on the order of twenty-five years, the compound annual rate of return in the high-income steady state is about 13.7 percent, and the amount of time devoted to training is about five and a half years. We could reduce the high-income steady-state rate of return, which is higher than most estimates of the postwar U.S. average, by choosing a lower value for λ, but this would mean a lower value for the amount of time devoted to training. If one views, say, high school education as part of the time devoted to training in modern economies, then five and a half years may already be too low. Our value of λ strikes a compromise on these competing aims.

We look to the artificial intelligence literature to set the parameters of the genetic algorithm. The minimal number of strings for effective search is usually taken to be 30. We chose to set $N = 50$ so that there are 50 agents *per generation* in our model, giving us a total population of 100 agents. We set the bit string length $\ell = 30$, with $\ell/2 = 15$ bits devoted to each of the decisions the agents face when young. String length is unimportant except as it determines the grid over which the agents can search for an optimal value. By setting the substring length to 15 bits, we effectively created a two-dimensional grid with $(32, 767)^2$ locations over a unit square and required the agents to choose optimal values on this grid. We set the probability of crossover, p^c, equal to .95, and we set the probability of mutation, p^m, equal to .0022. These values are close to those recommended by Grefenstette (1986). We now turn to the design of our computational experiments.

4.2. Experiment Set A: The Effects of Initial Conditions

4.2.1. Design of Experiment Set A. We first consider the effects of different initial conditions on the behavior of the system under learning. Our model has initial conditions for the per capita stock of human capital, x, the capital-to-effective labor ratio, k, the average initial fraction of time devoted to training, τ, and the average initial savings fraction, ϕ. We chose five feasible initial values for each of these four variables and simulated the system once for each possible combination of these five initial conditions. This yields 625 computational experiments, each with a different set of initial conditions. We conducted each experiment for 250 iterations and calculated the average of the last ten values of τ, denoted by $\bar{\tau}$, and the average of the last ten values of the capital-to-effective labor ratio k, denoted $\bar{k}$. Let us denote by k^* and τ^* the equilibrium levels of the capital-to-effective labor ratio and the training fraction at the two steady states. We examined the data to see if $|k^* - \bar{k}| < .002$ and $|\tau^* - \bar{\tau}| < .02$ at date t. If this criterion was met, we say that the system was in a neighborhood of that particular equilibrium at date t.

We chose the set of initial conditions as follows. Interesting initial x values are at or below $x|_{y(x)/\hat{y}=0.5} = .1$, the value of x that puts y at 5 percent of $\hat{y}$. We chose one initial $x(0)$ value higher than this and three lower; accordingly, we used five values of $x(0) \in (.0001, .001, .01, .1, 1)$. We set initial capital-to-effective labor ratios $k(0)$ relative to steady-state values according to $k(0) \in (.5k_\ell^*, k_\ell^*, .5k_\ell^* + .5k_h^*, k_h^* 1.5_h^*)$. Average initial τ and average initial savings fractions ϕ can range between zero and one. We chose five different values for each of these initial fractions in order to cover the whole range of possibilities: $\tau, \phi \in (.1, .3, .5, .7, .9)$. However, actual initial values for τ and ϕ are only approximately equal to our targeted values, due to the way in which we initialized strings.[14] We call this set of 625 computational experiments "experiment set A."

4.2.2. Results from Experiment Set A. One of the main results from experiment set A is that, depending on the settings of the initial conditions, neighborhoods of either of the two steady states can obtain at a point in time, which we limit to 250 iterations. Persistent mutation is the only source of variability in these neighborhoods. A typical time series from this set of 625 experiments reveals that the system initially fluctuates but then settles down to a neighborhood of either the low-income or the high-income steady state. These systems then remain in these neighborhoods for the remaining iterations. A sample time series is given in Figure 1, where it is a neighborhood of the low-income steady state that obtains.

A second key result from experiment set A is that, among the initial conditions, the initial level of human capital per capita, $x(0)$, is the dominant determinant of the behavior of the system at iteration 250. For low values of $x(0)$, we find the systems are in a neighborhood of the low-income steady state at iteration 250, while for high values of $x(0)$, we find the systems are in a neighborhood of the high-income steady state at iteration 250. Other initial conditions only influence this outcome for borderline values of $x(0)$. This result is interesting since preindustrial economies tend to be characterized by especially low levels of human capital per capita. Our model therefore predicts that these preindustrial economies will spend a long period of time, an *epoch*, in a neighborhood of the low-income steady state.

Figure 2 illustrates the importance of the initial level of $x(0)$. In each of the three sections, $k(0) = .001989$, but the results are the same for other values of $k(0)$. What varies in these three sections are the initial levels of $x(0)$. In Figure 2a, $x(0) = .001$; in Figure 2b, $x(0) = .01$; and in Figure 2c, $x(0) = .1$. In all three figures, the initial average fraction of wealth saved is plotted on the horizontal axis, and the initial average time devoted to training is plotted on the vertical axis. The actual initial average values for τ and ϕ are indicated by the placement of the labels *Low* or *High* in each of these figures. These labels,

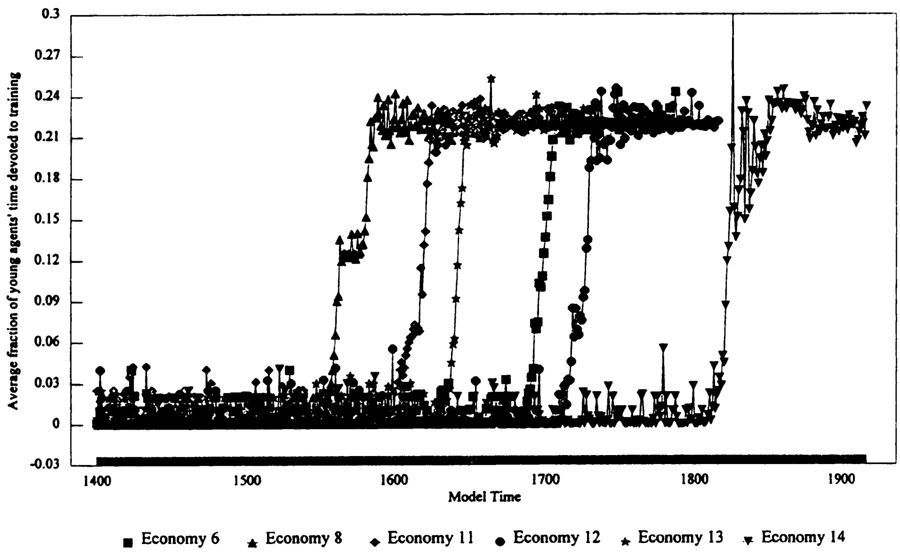

Figure 1. An epoch in a neighborhood of the low-income steady state.

Low and *High*, indicate whether our system had achieved neighborhoods around either the low- or high-income steady states of the model after 250 periods of model time. In Figure 2a, where $x(0) = .001$, the system is in a neighborhood of the low-income steady state after 250 iterations for all initial values for average τ and ϕ. For $x(0) = .01$, as illustrated in Figure 2b, the system has achieved the high-income steady state after 250 iterations in only three out of the twenty-five different combinations for initial average τ and ϕ. Notably, the three instances in which the system had achieved the high-income steady state were all cases for which the initial level of average τ was quite high to begin with (approximately .9), so that at least early on in the development process, there was a significant accumulation of human capital. This greater initial accumulation of human capital together with a higher initial stock of human capital $x(0) = .01$, perhaps along with some helpful mutations, enabled the system to achieve the high-income steady state. But Figure 2c demonstrates that this is simply a borderline situation. In Figure 2c $x(0) = .1$, and the system achieves a neighborhood of the high-income steady state after 250 iterations for all twenty-five combinations of initial average τ and ϕ. The only important difference between Figures 2a and 2c is the initial level of human capital per capita, $x(0)$. Thus, we see that the initial level of the stock of human capital plays a dominant role relative to other initial conditions. If $x(0)$ is relatively low, then we observe that the system is in a neighborhood of the low-income steady state at model time 250 regardless of other initial conditions, while if $x(0)$ is relatively high, we observe that the system is in a neighborhood of the high-income steady-state at model time 250, regardless of other initial conditions. This result holds across other values of $k(0)$, which is held constant in all three sections

Table 1. Results for Experiment Set A as a function of the initial level of human capital per capita.

Value of $x(0)$	Number of experiments	High steady state outcome at $t = 250$	Low steady state outcome at $t = 250$
.0001	125	0	125
.001	125	0	125
.01	125	18	107
.1	125	125	0
1.0	125	125	0

Note: Experiment set A consists of 625 experiments, one for each combination of initial conditions. The table lists results as a function of $x(0)$ only. For low values of $x(0)$, the low income steady state is observed at model time 250 regardless of other initial conditions.

of Figure 2. Further confirmation was obtained for the two other values for $x(0)$ that we considered, $x(0) = .0001$ and $x(0) = 1$. The case where $x(0) = .0001$ is qualitatively similar to the case where $x(0) = .001$, meaning that these 125 experiments were without exception in a neighborhood of the low-income steady state at iteration 250. Similarly, the case where $x(0) = 1$ is qualitatively similar to the case where $x(0) = .1$ because these 125 experiments were without exception in a neighborhood of the high-income steady state at iteration 250. Table 1 reports the results for experiment set A as a function of $x(0)$.

4.3. *Experiment Set B: Long-Run Behavior*

4.3.1. Eventual Attraction to the High-Income Steady State. While initial attraction to a neighborhood of the low-income steady state is likely for preindustrial economies—economies with low initial values for $x(0)$—both intuition and the results for experiment set B (given below) can be used to establish that these systems will eventually be attracted to a neighborhood of the high-income steady state *with probability 1*. The intuition is as follows. Suppose all agents have initially coordinated on the low-income steady state. The constant probability of mutation $p^m > 0$ implies that there will be some persistent experimentation by agents with nonzero investments in training—that is, there will sometimes be one or more agents who choose positive training amounts ($\tau_t^i(t) > 0$). How often such experimentation occurs depends, of course, on the mutation rate. This experimentation implies that effective labor units per unit of time worked (the average human capital that all agents inherit) will be rising over time. While the economy remains in the neighborhood of the low-income steady state, selection pressure will work against agents who invest positive amounts in human capital (training). The time they spend in training lowers the time they spend working, and the return from working more and investing more savings in physical capital dominates the return from investing in human capital at the low-income steady state. Decision rules that call for positive investments in training do not propagate and instead are systematically killed off. This keeps the system in a neighborhood of the low-income steady state for some time.

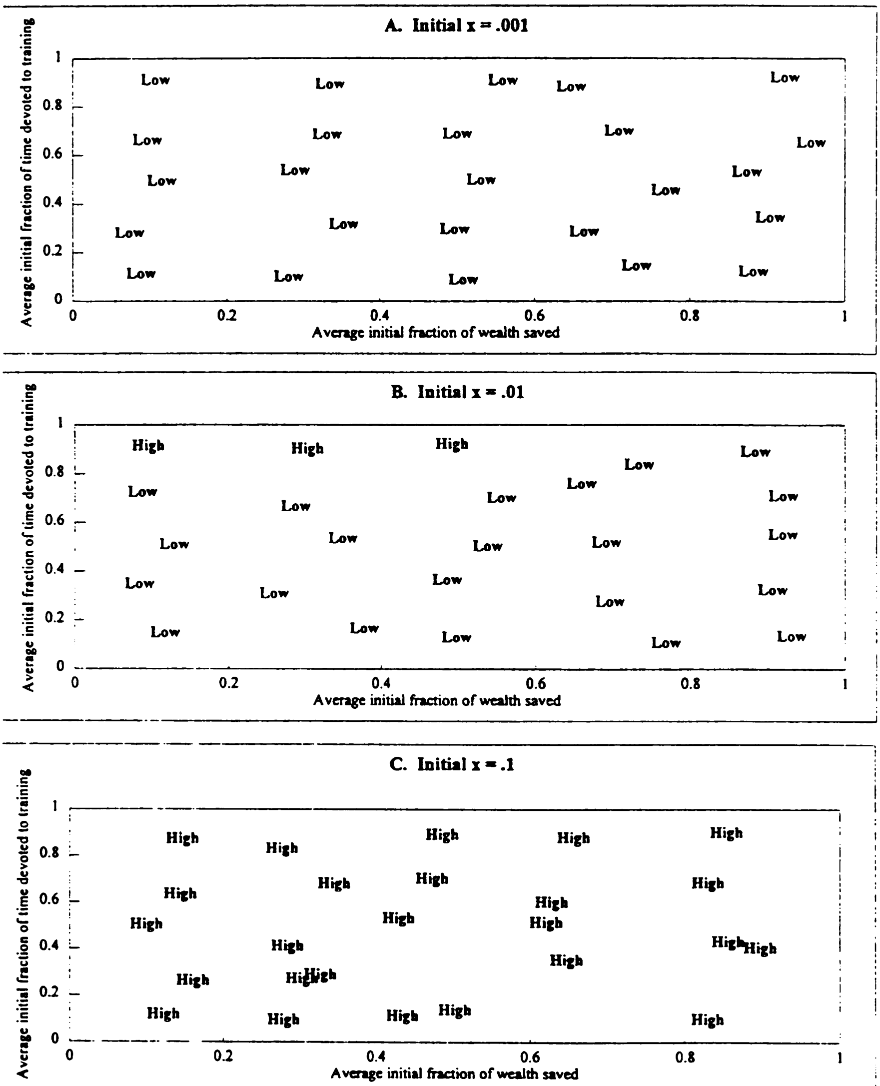

Figure 2. Behavior of economies at iteration 250 as a function of initial conditions.

However, since agents are always experimenting with positive amounts of training, the stock of human capital per capita, $x(t)$, grows slowly and unevenly according to the training technology and the law of motion for $x(t)$ until this stock of human capital eventually becomes large enough so that the rate of return to investing in human capital in equated with the rate of return to investing in physical capital. At this point, *selection pressure switches* because decision rules (strings) that call for investing positive amounts of time in training now yield higher fitness values than those decision rules (strings) that continue to instruct their owners to invest zero time in training. Thus strings that call for investing in training propagate, and the no-training strings are systematically killed off. Eventually, all agents are devoting a positive fraction of their youthful time endowment to training and this fraction lies within a small neighborhood of the fraction consistent with the high-income steady state. The economy stays in a neighborhood of the high-income steady state forever. Note that a transition, when it occurs, is always *from* a neighborhood of the low-income steady state *to* a neighborhood of the high-income steady state; a reversal from the high-income to the low-income steady state would never be optimal, and therefore such a reversal will not occur.[15]

A corollary to this intuition is that *initially identical* economies that spend an epoch in the neighborhood of the low-income steady state may have radically different dates of development takeoff. This occurs because the exact sequence of mutations that an economy experiences will determine which country reaches the threshold level of human capital first. While the date at which the transition occurs may well differ across countries we emphasize, once again, that the transition from a neighborhood of the low-income steady state to a neighborhood of the high-income steady state occurs with probability 1.

4.3.2. Design of Experiment Set B. In experiment set B, we verified the intuition provided in the previous subsection by studying the long-run behavior of these artificial economies computationally. We want to show that these economies always eventually attain the high-income steady state. We also want to study the timing of development takeoffs. To pursue these aims in the starkest possible way, we began each of fifteen computational experiments with exactly the same initial conditions and all parameters set identically, including the rate of mutation. The fraction of time devoted to training was zero for all agents, and the savings fraction was the one that is consistent with the low-income steady state for all agents. The value of $k(0)$ was the one that is consistent with the low-income steady state, and $x(0)$ was set to .01. The only difference between these computational experiments is that we used a different random number seed for each experiment. We terminated these experiments when our convergence criterion was met for the high-income steady state. For these experiments, our convergence criterion was to require $|k^* - \bar{k}| < .001$ and $|\tau^* - \bar{\tau}| < .001$, where $\bar{k}$ and $\bar{\tau}$ are calculated over the last thirty observations.[16]

4.3.3. Results from Experiment Set B. Our results from experiment set B verify the intuition given above, as all of the economies in this set of experiments initially remain in the neighborhood of the low-income steady state for hundreds of generations but eventually

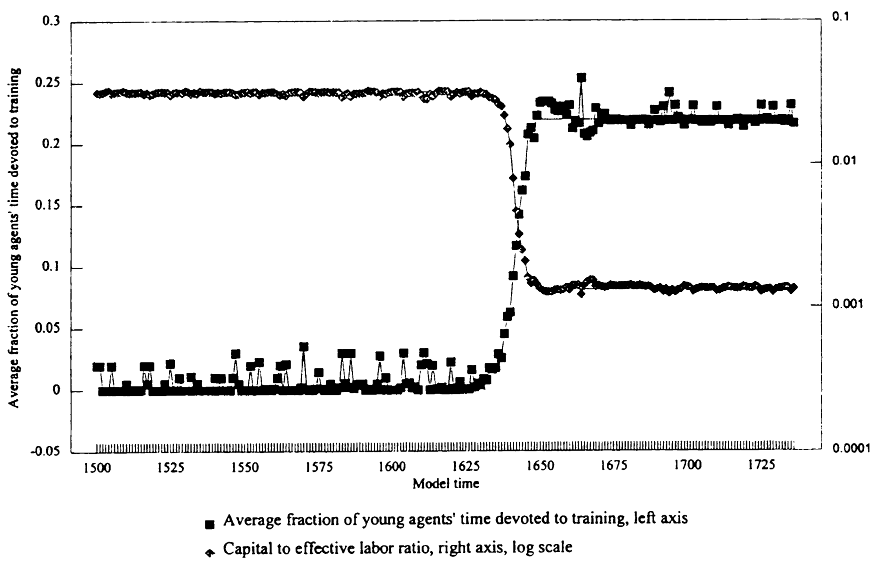

Figure 3. A development takeoff.

transit to a neighborhood of the high-income steady state. The results from experiment sets A and B constitute our claim that this model can address the fundamental fact of development and economic documented by Maddison (1982)—namely, that sixteen of today's most highly developed economies were once stagnant for centuries.

A time series of what occurs in a typical result from experiment set B is illustrated in Figure 3, which depicts a development takeoff. The average fraction of young agents' time devoted to training is measured on the left axis, while the capital-to-effective labor ratio is measured on the right axis. The low-income steady state values for k and τ are indicated by horizontal lines in the left portion of Figure 3, but they are difficult to see because the variables are close to steady-state values nearly all of the time. Agents have initially coordinated on a neighborhood of the low-income steady state (where $\tau = 0$) and remained there for the first 1,499 periods, which are not pictured. The economy remains in a neighborhood of the low-income steady state through model time 1,625 before it has, through experimentation, accumulated a sufficiently high stock of human capital. At this point, the rate of return to investments in human capital reaches that of the rate of return to investments in physical capital. A development takeoff occurs, and the population of artificial adaptive agents begins the process of adjusting their decisions for τ and ϕ accordingly. The economy transitions to a neighborhood of the high-income steady state, indicated by the two horizontal lines in the right half of Figure 3, where τ is now greater than zero. By about model time 1,675, the economy can be said to have coordinated on a neighborhood of the high-income steady state.

The remaining experiments in this set produced results qualitatively similar to those depicted in Figure 3. We checked at every iteration to determine whether our system had met our convergence criterion for the high-income steady state. The mean number of iterations at which our convergence criterion was met was 1,797 iterations, with a standard deviation of 70. The maximum number of iterations for convergence to the high-income steady state was 1,916 and the minimum number of iterations for convergence was 1,657. Even though all fifteen of these economies were initially identical and initially coordinated on a neighborhood of the low-income steady state, each nevertheless industrialized at a different time. If we interpret each generation as a period of roughly twenty-five years, the standard deviation of seventy iterations implies that a typical difference in the date at which the high-income steady state is achieved across societies according to these experiments is about 1,750 years. This figure is too large to apply directly to the international experience as we know it, but it does suggest that in this model *there is the possibility of a very wide disparity in the time it takes for countries to industrialize, even when countries all begin the process with exactly the same initial conditions.* We want to emphasize this feature as an interesting property of the model and caution against taking any particular calculation too seriously. The disparity in dates of industrialization could be reduced or increased, for instance, by either reducing or increasing the constant rate of mutation used or by reducing or increasing the value of $x(0)$.[17]

Figure 4 illustrates the different timing of the development takeoff for six of the fifteen artificial economies in experiment set B. We show only six economies in order to reduce clutter. This figure plots the average τ value in each of these six economies from iteration 1,399 through iteration 1,916, when the last of the fifteen artificial economies met our convergence criterion for the high-income steady state. All economies in experiment set B were in a neighborhood of the low-income steady state for the first 1,400 iterations. The development takeoff is illustrated as the movement of average τ from a neighborhood of the low-income steady state value, $\tau = 0$, to a neighborhood of the high-income steady state value, $\tau \cong .22$. Beginning at the low-income steady state, human capital per capita rises slowly and haphazardly across economies, since there is little private incentive to accumulate it. Because experimentation is a stochastic process, some economies reach the threshold level of human capital per capita before others; these countries industrialize rapidly and enjoy high growth subsequently. Other countries reach the threshold level of human capital per capita in due course but perhaps considerably later than the first group of countries. These countries then industrialize and eventually enjoy high growth, but their level of per capita income will be significantly lower than that of the countries that industrialized earlier and will remain lower even though the countries that industrialized later have achieved the high-income rate of growth. We can interpret the different timing of the development takeoffs that is illustrated in Figure 4 as being due to the different beliefs that agents have over time in the different economies about how much to invest in human and physical capital. The larger the amount of experimentation with nonzero investments in human capital, the faster a nation is able to reach the threshold level of human capital that is necessary for a development takeoff.

Differences in dates of industrialization can potentially go a long way toward explaining the differences in levels of per capita income across countries that we observe today. Consider a stylized calculation patterned after the model of this article. There are two steady

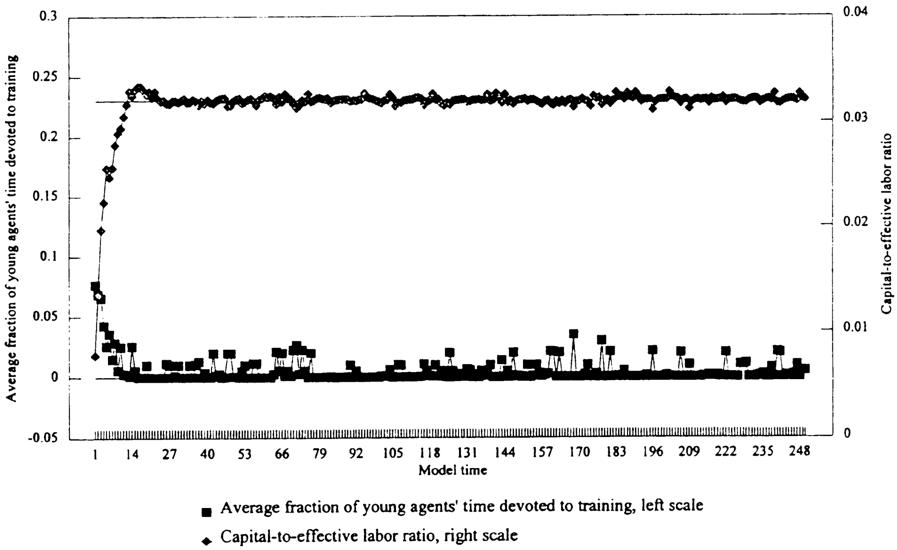

Figure 4. Six artificial economies industrialize.

states, one with no growth in per capita income and one with a growth rate of 1.6 percent per year. There are two countries, A and B, both initially in the low-growth equilibrium. Countries switch between steady states abruptly and without any transition time. Country A achieves the high-growth steady state in the year 1750, while country B achieves the high-growth steady state in 1960. If this is the situation, the ratio of per capita income in country A relative to country B in 1960 would be about 28.5. Both countries would grow at the same rate from 1960 through 1985, and so this ratio would remain constant. This is roughly consistent with the findings of Parente and Prescott (1993). This calculation is meant only to be illustrative, but we think it is suggestive that a two-steady-state model with learning providing a transition between the steady states can help address some of the main facts in economic development.

5. Remarks

Our modified version of an endogenous growth model is consistent with several broad development and growth facts. The modification we study is to introduce learning, which serves to select among equilibria and also provides a transition dynamic between stationary equilibria. We find that for low initial levels of human capital per capita—levels that tend to characterize preindustrial economies—and regardless of other initial conditions, the economies we study are initially attracted to the low-income steady state of the model and can remain there for long periods of time. Eventually, however, these economies achieve a high development state. These results are consistent with a fundamental development fact documented by Maddison (1982): today's leading industrialized nations were all growing at zero or near zero rates for centuries prior to the industrial revolution in Europe. Furthermore, in this model a development takeoff can occur at radically different times for two economies with identical initial conditions. These economies both eventually grow at the same mean rate, according to this model, but the level of per capita income will be significantly different in the two countries and will remain so indefinitely. This helps account for another fundamental development fact documented by Parente and Prescott (1993) and others: the level of per capita income is higher in the richest 5 percent of countries relative to the poorest 5 percent by a factor of 29, and furthermore, this factor has been constant from 1960 through 1985.

We chose to use a genetic algorithm to characterize the trial-and-error learning process of agents not only because this algorithm offers a natural model of experimentation by a heterogeneous population of agents, but also because the algorithm is well suited to "global" searches through its processing of many different decision rules simultaneously—in parallel. Indeed, the genetic algorithm, acting in concert with the training technology, is responsible for generating the transition from the low- to high-income steady state. Initially, for sufficiently low values of human capital per capita, the genetic algorithm achieves a neighborhood of the low-income steady state via the operations of reproduction, crossover, and mutation. The constant mutation rate of the algorithm ensures that there is constant experimentation with nonzero levels of time devoted to training, even though such positive levels of training yield relatively lower fitness values when all agents have coordinated on a neighborhood of the low-income steady state. However, this experimentation with positive amounts of training ensures that the stock of human capital rises over time, eventually achieving the threshold level. Once this threshold level is achieved, selection pressure changes as the genetic algorithm—acting once again via the operations of reproduction, crossover, and mutation—moves away from a neighborhood of the low-income steady state and achieves a neighborhood of the (now optimal) high-income steady state, remaining there indefinitely. We doubt that more mechanical (for example, recursive or gradient-based) learning models with noise could deliver these same results, though this remains, of course, an open question.

There are a number of possible extensions that could be made to the basic model that we have developed in this article. One extension would be to consider *neighborhood effects* (see, for example, Durlauf (1995))—that is, one could allow different, neighboring countries (different populations of artificial agents) to exchanges ideas (decision rules) about how much to save and how much to invest in human capital. If one nation had, for example, a greater propensity to experiment with human capital investments than another, the exchanges of ideas might have the effect of increasing the stock of human capital in the country with the lower propensity to experiment and thus speed up the development process in that country. Such neighborhood effects might explain, for example, why most of western Europe developed within the half century following the industrial revolution in Great Britain.

A second extension might be to include more than one threshold level for human capital accumulation. The purpose of this exercise would be to ascertain whether the country that was first to achieve the first threshold level for human capital (say, for example, Great Britain), would necessarily be the same country that was the first to achieve the second threshold level for human capital.

A final extension would be to consider how the distribution of income changes during the transition from the low- to the high-income steady state. We note that when agents have initially coordinated on the low-income steady state, income is distributed rather equally across agents. During the transition period to the high-income steady state there is rising income inequality as some agents do better (in terms of fitness) by investing more of their youthful time endowment in training rather than working. Once all agents have achieved a neighborhood of the high-income steady state, income once again becomes more equally distributed. The rising income inequality that we observe during the transition is largely consistent with Kuznets's (1955) hypothesis that the process of industrialization leads to a rise in income inequality and that it is only at relatively high levels of income that such inequality is reversed.[18] We leave these extensions to future research.

Appendix: Program Outline

Begin program.
    Set model parameters.
    Set simulation parameters.
    Set genetic algorithm parameters.
    Find equilibria numerically: (τ_ℓ^*, k_ℓ^*) and (τ_h^*, k_h^*).
    Initialize tensors and accumulation variables.
    Initialize k, x.
    For replications=1 to maximum replications,
        Initialize strings;
        Find implied values of y, w, r for $t = -2$;
        Find implied initial old aggregate savings;
        Find implies values of k, y, w, r, x for $t = -1$;
        Find implied initial young aggregate savings;
        Find implied values of k, y, w, r, x for $t = 0$;
        For time=1 to maximum time,
            Find fitness of the old generation;
            Find fitness of the young generation;
            Create newborn generation: for member=1 to generation size;
                Apply reproduction operator via tournament selection;
                Apply crossover and mutation operators;
            End creation loop;
            Find aggregate savings of newborn generation;
            Find values of k, y, w, r, x for time t;
            Delete old strings, let old=young, let young=newborn;
        End time loop;
    End replication loop.
    End program.

Acknowledgments

The authors thank Oded Galor, Robert Tamura, and an anonymous referee for helpful comments and suggestions. The views expressed in this article are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

Notes

1. The sample consists of Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Sweden, Switzerland, the United Kingdom, and the United States.
2. See, for example, Murphy, Shleifer, and Vishny (1989), Becker, Murphy, and Tamura (1990), Azariadis and Drazen (1990), Matsuyama (1991), and Laitner (1995) among others.
3. See Azariadis (1996) for an introduction to the economics of poverty traps.
4. An exception is Galor and Tsiddon (1997), which is discussed below.
5. We could have allowed for more than two steady states, but we elected to study a stylized two-steady-state case in this article in order to discuss the main ideas in the clearest possible way.
6. For an introduction to genetic algorithms, see Goldberg (1989) or Michalewicz (1994).
7. For an alternative approach to systems with heterogeneous learning rules, see Evans, Honkapohja, and Marimon (1995).
8. In this article, we use the term *global* to describe our analysis because it is based on a random initialization scheme. We recognize that our analysis is not truly global, even computationally speaking, since we did not complete multiple experiments based on every possible initialization for a given parameterization. Such an approach is beyond the scope of this article.
9. We could include exogenous labor-augmenting technological change and population growth, but these factors would exogenously increase the output growth rate in both steady states and serve only to complicate the analysis. For this reason we follow Azariadis and Drazen (1990) and abstract from these factors by assuming a constant population and a static technology.
10. The assumption that capital depreciates fully each period is not necessary to our results; it merely simplifies our analysis.
11. We add one so that neither fraction can be equal to unity. If either $\tau_t^i(t)$ or $\phi_t^i(t)$ is equal to one, the consumption $c_t^i(t)$ for that agent is zero, implying utility of $-\infty$. This causes a computational problem which we avoid by using 32,768 instead of 32,767. This also explains why we restrict τ and ϕ to be chosen from the interval $[0, 1)$ rather than from the interval $[0, 1]$.
12. Of course, some strings *were* in use in the previous period, so this question might seem a little redundant. We phrase the question this way only to emphasize that we ask the same question of every string in the system at time t in order to evaluate all strings on an equal basis.
13. The addition of crossover serves to speed convergence somewhat, but it is the constant mutation rate $p^m > 0$ rather than the crossover operation that is responsible for our main results. We note that while crossover may serve to speed convergence, it has little effect when the economy is in the neighborhood of an equilibrium; in this case, strings are already nearly identical, and so crossover plays almost no role in altering strings.
14. Our initialization procedure worked as follows. If we wanted to initialize τ and ϕ so they were, say, both equal to .1, we would choose each bit value in each string by first choosing a random number from .01 to 1.00; if the random number was less than or equal to .1, we would place a bit value of 1 in that spot in the string, otherwise, we would place a bit value of 0. Since we only have fifty agents in each young generation, our initial values for τ and ϕ are only approximately equal to our targeted initial values of .1, .3, .5, .7, and .9. This approximation is not material to our results.

15. A reversal from the high- to the low-income steady state cannot occur in our model because human capital only accumulates. Thus, the private return to investing in human capital is nondecreasing over time. Even a wildly improbable round of mutations that set all agent's training decisions to zero could not be sustained as a return to the low-income steady state because the private return to investing in human capital would remain sufficiently high (that is, above the threshold). This is a fundamental difference between our model and that of Kandori, Mailath, and Rob (1993).
16. We limited the number of experiments in this set to fifteen mainly to conserve on computation time. The qualitative results were unchanged in a number of other computational experiments that we did not organize into a reportable format.
17. Perhaps more important, we are following Azariadis and Drazen (1990) in abstracting from the possibility that labor or ideas or both can move across economies. We expect that a model including some degree of human capital mobility would mitigate the sharp disparities in dates of industrialization that we find. From this point of view, the results we obtain are perhaps reasonable for a world of completely isolated societies.
18. See also Galor and Tsiddon (1997), who obtain a similar result.

References

Andreoni, J., and J.H. Miller. (1995). "Auctions with Artificial Adaptive Agents," *Games and Economic Behavior* 10, 39–64.

Arifovic, J. (1994). "Genetic Algorithm Learning and the Cobweb Mode." *Journal of Economic Dynamics and Control* 18, 3–28.

Arifovic, J. (1995). "Genetic Algorithms and Inflationary Economies," *Journal of Monetary Economics* 36, 219–243.

Arifovic, J. (1996). "The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies," *Journal of Political Economy* 104, 510–541.

Arthur, W.B., J.H. Holland, B. LeBaron, R. Palmer, and P. Taylor. (1996). "Asset Pricing Under Endogenous Expectations in an Artificial Stock Market," SSRI Working Paper 9625, University of Wisconsin.

Azariadis, C. (1996). "The Economics of Poverty Traps Part One: Complete Markets," *Journal of Economic Growth* 1, 449–486.

Azariadis, C., and A. Drazen. (1990). "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105, 501–526.

Becker, G.S., K. Murphy, and R. Tamura. (1990). "Human Capital, Fertility and Economic Growth," *Journal of Political Economy* 98, S12–S37.

Binmore, K.G., and L. Samuelson. (1992). "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory* 57, 278–305.

Bullard, J., and J. Duffy. (1996a). "A Model of Learning and Emulation with Artificial Adaptive Agents," Forthcoming, *Journal of Economic Dynamics and Control.*

Bullard, J., and J. Duffy. (1996b). "On Learning and the Stability of Cycles," Forthcoming, *Macroeconomic Dynamics.*

Durlauf, S.N. (1995). "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," in W.A. Barnett et al., (eds.), *Disequilibrium Dynamics: Theory and Applications*. Cambridge: Cambridge University Press.

Durlauf, S.N., and P.A. Johnson. (1995). "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics* 10, 365–384.

Evans, G.W., S. Honkapohja, and R. Marimon. (1995). "Global Convergence in Monetary Models with Heterogeneous Learning Rules," Working Paper, University of Edinburgh, University of Helsinki, and Universitat Pompeu Fabra.

Galor, O., and D. Tsiddon. (1997). "The Distribution of Human Capital and Economic Growth," *Journal of Economic Growth* 2, 93–124.

Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley.

Goodfriend, M., and J. McDermott. (1995). "Early Development," *American Economic Review* 85, 116–133.

Grefenstette, J.J. (1986). "Optimization of Control Parameters for Genetic Algorithms," *IEEE Transactions on Systems, Man, and Cybernetics* 16, 122–128.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* Ann Arbor: University of Michigan Press. Also Cambridge, MA: MIT Press, 1992.

Holland, J.H., and J.H. Miller. (1991). "Artificial Adaptive Agents in Economic Theory," *American Economic Review* 81, 365–370.

Kandori, M., G.J. Mailath, and R. Rob. (1993). "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* 61, 29–56.

Kuznets, S. (1955). "Economic Growth and Income Inequality," *American Economic Review* 49, 1–28.

Laitner, J. (1995). "Structural Change and Economic Growth," Working Paper, University of Michigan.

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

Maddison, A. (1982). *Phases of Capitalist Development.* New York: Oxford University Press.

Marimon, R., E. McGrattan, and T.J. Sargent. (1990). "Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents," *Journal of Economic Dynamics and Control* 14, 329–373.

Matsuyama, K. (1991). "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics* 106, 617–650.

Michalewicz, Z. (1994). *Genetic Algorithms + Data Structures = Evolution Programs* (2nd ed.). New York: Springer-Verlag.

Miller, J.H. (1996). "The Coevolution of Automata in the Repeated Prisoner's Dilemma," *Journal of Economic Behavior and Organization* 29, 87–112.

Murphy, K., A. Shleifer, and R. Vishny. (1989). "Industrialization and the Big Push," *Journal of Political Economy* 97, 1003–1026.

Parente, S.L., and E.C. Prescott. (1993). "Changes in the Wealth of Nations," *Quarterly Review* (Federal Reserve Bank of Minneapolis) (Spring), 3–16.

Quah, D. (1996a). "Convergence Empirics Across Economies with (Some) Capital Mobility," *Journal of Economic Growth* 1, 95–124.

Quah, D. (1996b). "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics," *Economic Journal* 106, 1045–1055.

Romer, P.M. (1986). "Increasing Returns and Long-Run Growth," *Journal of Political Economy* 94, 1002–1037.

Routledge, B. (1994). "Artificial Selection: Genetic Algorithms and Learning in a Rational Expectations Model," Working Paper, University of British Columbia.

Sargent, T.J. (1993). *Bounded Rationality in Macroeconomics.* New York: Oxford University Press.

Summers, R., and A. Heston. (1991). "The Penn World Tables (Mark 5): An Expanded Set of International Comparisons, 1950–1988," *Quarterly Journal of Economics* 106, 327–368.

Tesfatsion, L. (1995). "A Trade Network Game with Endogenous Partner Selection," Economic Report Series No. 36, Iowa State University.

Wright, R. (1995). "Search, Evolution and Money," *Journal of Economic Dynamics and Control* 19, 181–206.