19

Sources of Technological Divergence Between Developed and Less Developed Economies

Raaj Kumar Sah and Joseph E. Stiglitz

Developed and less developed countries differ in numerous ways. Incomes per head are different, capital stocks are different, the levels of education of their citizens differ. Whether a cause or consequence of these differences, some of the most salient differences are related to technological change and innovation. Firms in many developed countries appear to be constantly looking around for new technologies, for better and more efficient ways of doing things; firms in the newly industrialized countries (NICs) and those LDCs aspiring to become NICs appear to be searching among available technologies for those which will most improve their productivity, while firms in the less successful LDCs are often slow to abandon the traditional ways of doing things. Though exceptions to this polar characterization can be found, the fact remains that the technological gap between many poorer LDCs and the developed countries has not significantly narrowed over the past 50 years.

Indeed, this constitutes one of the central puzzles facing development economics. Why is it that the growth rates and income levels of various countries have not converged faster than they have? There is some evidence (Baumol, 1986) that for a significant number of countries, there has in fact been divergence rather than convergence. Traditional neoclassical growth theory (Solow, 1956) predicts that in the long run, the growth rates in all countries should be related only to the rate of technological progress and the rate of population growth; growth rates in per capita income should be related only to the rate of labor-augmenting technological progress, and differences in levels of per capita consumption should be related to differences in savings rates.[1]

In this paper, we describe several different perspectives on the sources of non-convergence. Our emphasis is on one important aspect common to all of these perspectives, namely, that an economy can have multiple *equilibria*. That is, some societies may be characterized by high levels of innovation and others by low levels of innovation. One perspective is based on certain characteristics of technology. Some aspects of this approach have been pursued elsewhere by Stiglitz (1987), Lucas (1985) and others. The other is based on socioeconomic considerations, and this paper will explore these.

The objective of this paper is to provide an overview of these alternative perspectives and to sketch some simple models illustrating the points at issue, rather than to provide a detailed development of any of these approaches.

The paper is organized as follows: in section 1 we discuss the sources of multiplicity of equilibria in general terms. In section 2 we describe an approach which emphasizes individuals' learning about certain aspects of the economic environment. These models of social osmosis have been developed by Sah (1985, 1987a, 1988a) to analyze patterns of crime, corruption and other phenomena. Section 3 provides a general discussion of the sources of positive feedback which give rise to a multiplicity of innovative equilibria. Section 4 provides a brief discussion of several other perspectives which deal with the role of individuals' preferences (tastes) and with the birth and death of organizations. Section 5 presents certain aspects of the technological perspective.

This paper is dedicated to the memory of Carlos Díaz-Alejandro.[2] He inspired our attempt to broaden the perspectives with which we look at economics, particularly the problems of economic development. He had the rare ability to combine the techniques of modern economics with the insights and vision of an economic historian. He was not confined by the straightjacket imposed by the conventional paradigms: he knew those paradigms well, and he knew their strengths and weaknesses. He was a social scientist: his objective was to understand the development process in all of its ramifications, and to do this, he was willing to use whatever techniques and viewpoints that he thought helped obtain insights. We think he would have liked what we have attempted here. We would have vastly profited from his reactions to our analysis. We miss him.

1. Outline of the Theory

In this paper, we wish to put forward a theory of social equilibrium, in which it will turn out that in equilibrium some societies may be characterized by high levels of innovation, and others by low levels. (This view is sharply different from the kind of technological determinism popular in the late eighteenth and nineteenth centuries.) A central hypothesis underlying our analysis is that a major determinant of the survival value of any characteristic in the population – as well as a major determinant of the behavior of individuals in any society or organization – is the nature of the environment. But economic environments are endogenous. The nature of the environment should be explained by the theory. Of course, a central aspect of that environment is the mix of characteristics of those who make up the population. Thus, the mix of characteristics and behaviors of individuals in the environment is central in determining the kinds of characteristics which have survival value and in determining observed behaviors, the kinds of actions which individuals undertake. The mix of characteristics and behaviors at one date determines the set of characteristics and behaviors at later dates.

Specifically, we argue that society may be characterized by a high level of innovativeness, or by a low level of innovativeness. We explore a number of mechanisms through which the environment affects individuals' decisions to be innovative, and through which individuals' decisions to be innovative affect the socioeconomic environment. Several of these mechanisms can be expressed in conventional supply and demand terms: the demand for 'innovativeness' is an increasing function of the fraction of the population which is innovative; and the supply of innovativeness at any time is an increasing function of the fraction of individuals who are innovative in the previous period and of how well those innovative individuals have done. As a result there may exist multiple equilibria, in one (or several) of which there is a relatively small fraction of individuals who are innovative, while in others this fraction may be relatively large. Which of the multiple equilibria an economy exhibits at any particular time is partly a consequence of the historical path the economy has taken. Thus, the rate of technological progress in a society may depend explicitly on its history.

1.1 Feedbacks and Multiplicity of Equilibria

The principal source of multiplicity of equilibria here (as in other economic models) is positive feedback. To see this most clearly, consider the traditional economic model of a single market. Such models stress the role of negative feedbacks. An increase in the demand for some commodity results in an increase in price, and the increase in price brings forth an increase in supply, which dampens the initial price response. We draw the demand curve as downward sloping, the supply curve as upward sloping, and there is a unique equilibrium. Thus, if p is the price, $D(p)$ is the demand, $S(p)$ is the supply and $e(p)$ is the net demand, then the equilibrium condition

$$e(p) = D(p) - S(p) = 0 \tag{1}$$

yields a unique equilibrium if the derivatives D_p and S_p are respectively negative and positive.

We know, of course, that it is easy to generate analogous models with multiple equilibria. If the demand for a commodity increases the demand for a factor, and if the preferences of the owners of this factor are relatively intensive in that commodity, then there may be multiple equilibria. In one equilibrium, the factor has a high income because the demand for the commodity which uses that factor is high, and the demand for that commodity is high because the factor income is high. In another equilibrium the factor income is low because the demand for the commodity using that factor intensively is low, and the demand for that commodity is low because the factor income is low.[3] Still, by and large, economists believe that such effects are not too large, and that our conventional demand and supply stories, with their unique intersection, provide a good description of most markets.

This intuition, however, becomes much less reliable when we consider aspects of the economic environment in which prices do not play a central and direct role. The actions of one individual or group of individuals may have effects on others which are not mediated through the price system. These effects are like externalities. The nature of these interactions frequently gives rise to multiple equilibria; indeed, these provided the basis of the traditional Big-Push theories of development (Rosenstein-Rodan, 1943).[4]

1.2 The Screening Model: An Example

An illustration of non-price interactions giving rise to multiple equilibria is provided by the screening model of Stiglitz (1975, 1984). There are two types of individuals, the more able, with productivity γ_1, and the less able, with productivity γ_2. A fraction m of the population is more able. The more able can obtain certification of their ability, at a cost of c. If they do, they receive a wage corresponding to their ability. If a fraction θ of the more able get certified, then the average productivity (and the wage) of those who do not get certified is

$$W(\theta) = [(1-\theta)m\gamma_1 + (1-m)\gamma_2]/[(1-\theta)m + (1-m)] \tag{2}$$

Hence, the net gain from screening for someone with screening costs c, when the fraction of the able population obtaining the credential is θ, is

$$\gamma_1 - w(\theta) - c \tag{3}$$

The equilibrium is characterized by

$$\gamma_1 - W(\theta) = c \tag{4}$$

with corner solutions at

$$\theta = 0 \text{ if } \gamma_1 - w(0) < c$$

$$\theta = 1 \text{ if } \gamma_1 - w(1) > c$$

As more individuals obtain screening, the wage of the unscreened is reduced, so that the return to screening is increased. It is easy to see, as in figure 19.1, that there may exist multiple equilibria.

In the simplest of such models, with *only* positive feedbacks and identical individuals, equilibria naturally entail corner solutions in which either everyone or no one takes some action (in Stiglitz' model, obtaining a particular education level).[5] It is easy, however, to construct multiple interior equilibria by allowing both positive and negative feedbacks or differences among individuals.

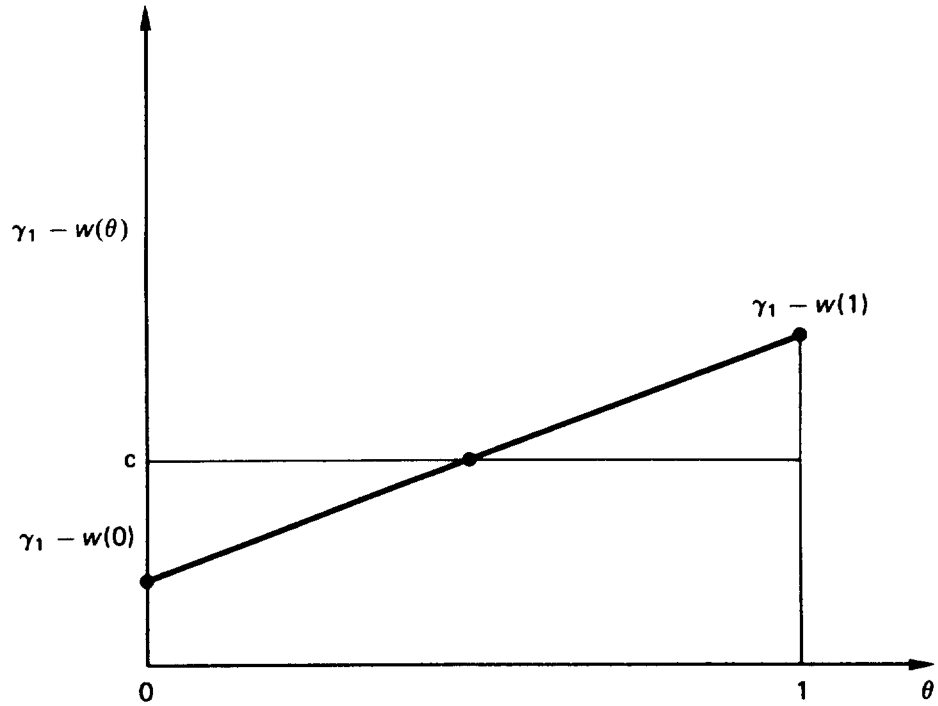

Figure 19.1 Multiple equilibria with corner solutions in screening model

1.2.1 Multiple Equilibria with Individual Differences Assume there is a distribution of costs of certification, $F(c)$. Then those with the lowest certification costs will get certified. As more individuals get certified, the benefits of certification increase but the costs of the marginal individual becoming certified also increase. Equilibrium is characterized by

$$F[\gamma_1 - W(\theta)] = \theta \tag{5}$$

Figure 19.2 shows the possibility of multiple interior equilibria.[6]

Related models are also useful in a variety of circumstances. Stiglitz (1974) has used such a model as a basis for multiple equilibria involving discrimination, where one group with some easily identifiable characteristic is caught in one equilibrium, say with a low level of education, and another group is in another equilibrium, say with a high level of education. Employers pay wages which reflect the productivity of the workers, given the information which is available to them, and workers in each group choose their optimal level of education, given the wages being paid by firms to members of their group. Starrett (1976) has analyzed alternative economic explanations of some of the radical views concerning income distribution.[7]

1.2.2 Multiple Equilibria with Positive and Negative Feedbacks Negative feedbacks arise naturally in many economic models and, in conjunction with positive feedbacks, give rise to multiple equilibria. In our example, if those who are identified as the more able work in different jobs than the less able

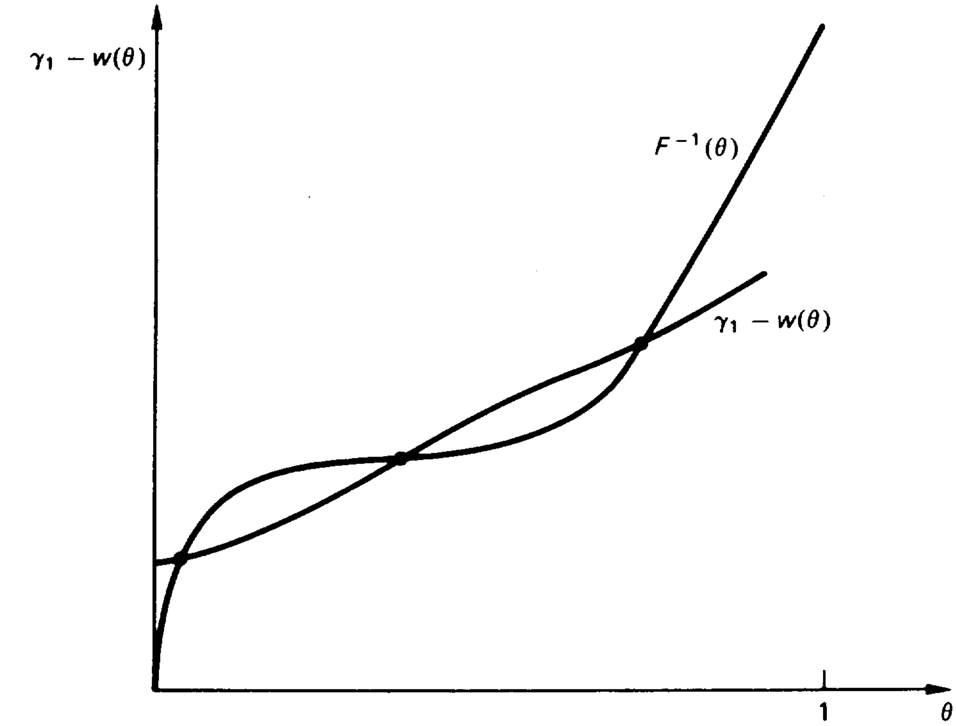

Figure 19.2 Multiple equilibria with individuals differing in costs of screening. $F^{-1}(\theta)$ gives the cost of screening at the θth percentile of the population

(the unskilled), then as more able individuals are identified, diminishing returns set in. Assume, for instance, that there is a production function of the form

$$Q = G(E_s, E_u)$$

where Q is output, E_s is effective skilled labor supply (number of individuals assigned to skilled jobs, multiplied by their average productivity) and E_u is the effective unskilled labor supply. Assume G exhibits constant returns to scale. We assume that the type 2 (unskilled) workers have zero productivity when assigned to the skilled job, but type 1 (skilled) workers are still more productive than type 2 workers when assigned to unskilled jobs. With appropriate normalizations, we denote the 'effective' units of labor of a skilled individual assigned to a skilled job as 1, of a type 1 worker assigned to an unskilled job as γ_1 and of a type 2 worker assigned to an unskilled job as γ_2. If the total population is L, and the fraction of type 1 workers screened is θ, then if all of the able who are identified as such are assigned to skilled jobs

$$E_u = [(1 - \theta)m\gamma_1 + (1 - m)\gamma_2]L$$

and

$$E_s = \theta mL$$

Let $e = E_u \backslash E_s$; let $W_s(\theta)$ and $W_u(\theta)$ be the skilled and unskilled wage, respectively, when a fraction θ of the able are screened; and let $g = G/E_s$ be output per effective unit of skilled labor, which, by the assumption of constant returns to scale, is just a function of e. Then

$$W_s = g(e) - eg'$$

$$W_u = g'(e)h$$

where h is a measure of the average productivity of the unscreened:

$$h = [(1 - \theta)m\gamma_1 + (1 - m)\gamma_2]/[(1 - \theta)m + (1 - m)]$$

In competitive equilibrium, all of the more able who are so identified will be assigned to skilled jobs, so long as

$$W_s > g'(e)\gamma_1 = W_u\gamma_1/h$$

and we focus attention on that situation here.[8] Equilibrium then entails

$$W_s - W_u = c$$

Differentiating with respect to θ, we obtain

$$\mathrm{d}(W_s - W_u)/\mathrm{d}\theta = - [g'' E_u/L^2\theta m(1-\theta m)]\ \mathrm{d}e/\mathrm{d}\theta - g'\mathrm{d}h/\mathrm{d}\theta$$

where $g' < 0$, $\mathrm{d}e/\mathrm{d}\theta < 0$ and $\mathrm{d}h/\mathrm{d}\theta < 0$. The second term is positive (the positive feedback effect that we identified earlier), while the first term is negative, as a result of diminishing returns. As figure 19.3 illustrates, it is easy to construct examples in which there are multiple interior equilibria, with varying proportions of the more able identifying themselves. In each, the *difference* between the wage of the skilled and the wage of the unskilled

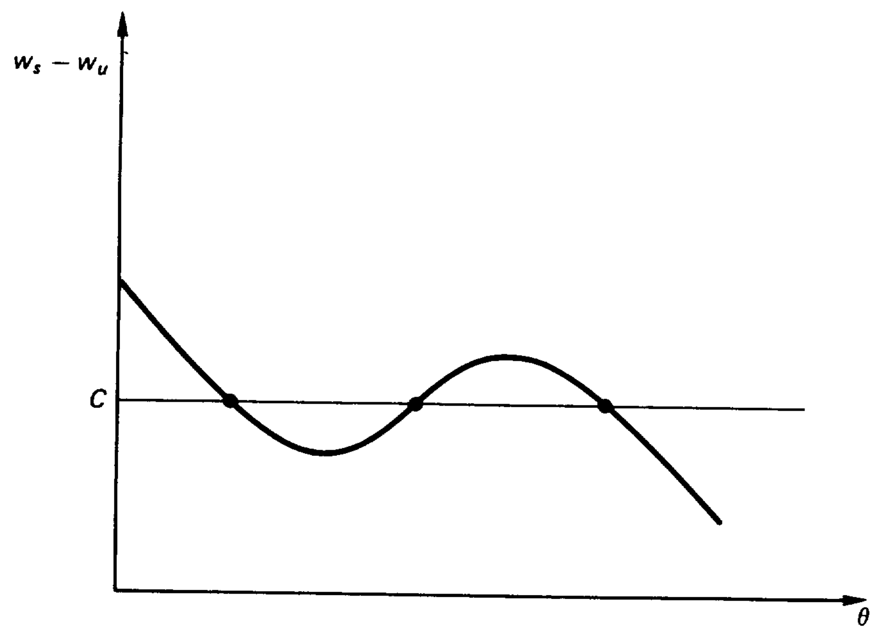

Figure 19.3 Multiple interior equilibria with fixed wage differential is fixed, but as more individuals are screened, the skilled wage falls; hence, the equilibrium in which the fewest able are screened Pareto dominates the other equilibria.

1.3 Innovation and Multiple Equilibria

In the discussion below, we consider several alternative reasons why there may be important positive feedbacks associated with having more innovation (more innovative individuals) in an economy, resulting in a multiplicity of equilibria, with some equilibria characterized by low levels of innovation, others by a higher level. Several of these models involve multiple equilibria associated with differences in individuals arising from differences in information about the environment or about the returns to being innovative, from differences in their characteristics which affect the 'costs' of being innovative or from differences in tastes.

1.4 Dynamics

Most of this paper focuses on steady-state (long-run) equilibria. But, associated with each of the models there are dynamics which describe the evolution of the economy over time and which determine whether any particular equilibrium is stable or unstable. The dynamics are interesting in their own right. Because the dynamics associated with each of the different perspectives we explore in this paper differ, the analysis of the dynamics may also provide suggestions for discriminating among theories. Understanding dynamics may also be important for identifying policy interventions to move the economy successfully from one equilibrium to another.[9]

1.4.1 Dynamics in the Screening Model Some of the central dynamic issues are again illustrated by our screening model. Consider the version presented above where individuals differ in their costs of being certified. Individuals base their decisions about education (certification) on *expectations* concerning the wages they will receive if they are, or are not, certified. The simplest assumption is that expectations are based on wages *currently* being received by those who are certified and those who are not. This is the simplest extrapolative model. Firms are assumed to adjust wages quickly (relative to the time it takes to change education norms) to reflect true mean values of marginal productivities. In this view, the dynamics are given by

$$F[\gamma_1 - W(\theta_t)] = \theta_{t+1}$$

Let θ^* be an equilibrium that is a solution to (5). Equilibria in which, for values of θ slightly less than θ^*, equilibrium returns to being screened exceed the screening costs of the marginal individual (and conversely for values of θ slightly greater than θ^* in equilibrium) are stable, while other equilibria are unstable. Thus, in figure 19.2 the first and third equilibria are stable.

For this model to be a good description, individuals do not have to know the actual value of $\theta(t)$. What is important is that they know (or have beliefs about) the returns to being certified, which are a function of $\theta(t)$.

One could argue about the reasonableness of our expectations hypothesis on the grounds that it is too simplistic, particularly in the simple environment which we have postulated. Individuals should learn that $\theta(t)$ is not equal to $\theta(t-1)$. If they learn to correctly extrapolate, this would, of course, alter their behavior. If they come to believe that all others are also learning the process by which decisions are made, then the only possible equilibria are the Nash equilibria, the solutions to equation (5). More generally, it should be apparent that the nature of dynamics and the resulting equilibria depend on the expectation formation processes. We discuss some limited processes below.[10] The next section describes individuals' learning and its consequences in certain circumstances in which individuals have limited information and knowledge. In section 3 we discuss some effects of other information-generating mechanisms. In section 6.2 we argue why the introspection-based predictions emphasized in many strategic models may not be an appropriate description of many complex situations in which individuals find themselves.

2. Individuals' Learning about some Aspects of the Environment

In this section, we emphasize the role of individuals' learning about certain relatively intangible economic aspects of their environment which might be important in their decision to innovate. For concreteness, we begin by describing some elements of the analysis Sah (1985, 1987a, 1988a) has developed to explain certain patterns of phenomena such as crime, corruption and honesty. The assumptions underlying an individual's behavior are as follows:

1 As part of his decision making, an individual needs to predict some environmental variables about which public or market-based information is unavailable. For instance, in an individual's choice to undertake an illegal tax transaction, he needs to estimate the probabilities of encountering a corrupt versus non-corrupt tax bureaucrat. The difficulties of ascertaining this probability are obvious; not surprisingly, in most countries there are no statistics – let alone good statistics – on the proportion of corrupt tax bureaucrats.
2 The relevant information on which an individual bases his estimates and decisions is highly limited. It consists of his own past experiences, and it may also partly reflect observations gathered from others. The information is quite obviously limited, partly because there are increasing costs of acquiring more personal or indirect observations. Indirect observations further entail costs of miscommunication.
3 There is a significant stochastic element to any individual's observations,

and hence to his information and beliefs, and the observations are affected by actual past variables. Thus, if a larger fraction of tax bureaucrats were actually corrupt in a past period of an individual's life, then his observations concerning that period will partly, and stochastically, reflect this fact.

4 The emphasis here is on those types of individual decisions in which strategic considerations are not central. For instance, since an individual can encounter any of a large number of tax bureaucrats, it is reasonable to assume that the individual does not believe that his actions will significantly affect actions undertaken by others. (We discuss strategic considerations later in the paper.)

To see some of the consequences of this type of analysis in the context of technological change, consider a stylized model in which an individual chooses to be either an innovator or a non-innovator ('bureaucrat') in each period.[11] Define a variable $x^h(t, \lambda)$ such that $x^h(t, \lambda) = 1$ if a person with characteristics h, who has lived for λ periods, is an innovator in period t. $x^h(t, \lambda)$ is zero otherwise. Let $y(t)$ denote the number of innovators in period t. Thus

$$y(t) = \sum_{\lambda=0}^{L-1} \sum_{h} x^h(t, \lambda) \tag{6}$$

where L represents the life span of an individual.

For simplicity, we assume that in any period, the individual interacts with one other individual, and his pay-off depends both on what he does (i.e., whether he chooses to be an innovator) and on what the other individual has done. For an individual who chooses to be an innovator in a period, let U^{11} be the utility if he encounters an innovator, and let U^{12} be his utility if he encounters a bureaucrat. For an individual who chooses to be a bureaucrat in a period, let U^{21} and U^{22} denote the corresponding utilities. Assume for the moment that[12] $U^{11} > U^{21}$ and $U^{22} > U^{12}$. That is, being an innovator is better if one were to encounter an innovator, and being a bureaucrat is better otherwise. Then, the assumptions outlined earlier imply that the probability that an individual will choose to be an innovator will be influenced by the number of innovators in some of the past periods. That is, $Pr[x^h(t, \lambda) = 1] = g[y(t -), \ldots, y(t - \lambda)]$. From equation (6) therefore one obtains a relationship

$$y(t) = f[y(t-1), \ldots, y(t-L+1)] \tag{7}$$

This relationship has the property that the current number of innovators in the population will be higher if the number of innovators has been higher in the past, and that the economy can have multiple steady-states.[13]

3. Sources of Positive Feedback in Innovation

Recall the stylized assumption made above that it is better for an individual to be an innovator if he were to encounter an innovator. Such a reward structure can arise from a variety of sources. The response a worker gets to a new idea proposed to his boss depends on whether his boss is 'innovative' (or acts innovatively) or not.

In other circumstances interactions involve several individuals. Most changes in methods of production require the acquiescence of many individuals. Changes in production are, at least to some extent, 'political' decisions. An owner who finds a better way of producing a widget must convince his banker to lend him money. Another dimension, which is perhaps far more important in LDCs than in developed countries, is that an owner or a manager also needs to obtain approval from a variety of government officials.

Positive feedbacks such as these also arise indirectly, through the negative effects of non-innovative individuals on innovators, and conversely, through the negative effects of innovators on non-innovators. A bureaucrat may have the power to suppress an innovation, or at least to make it more difficult for an innovation to occur. Innovative individuals on the other hand can create situations where the routines loved by bureaucrats are simply unproductive. In all of these cases, the return to being innovative depends (positively) on the fraction of those with whom one must interact that are innovative; and the individual's beliefs about this depend on past experience.

In still other cases, the return to being innovative depends more directly on the number of innovative individuals in the economy.[14] It may depend positively on the number of those in other industries who are engaged in innovation, for often ideas which prove useful in one context can be transferred to another. Moreover, the return to the inventors, to those who generate new ideas, depends in part on how likely it is that their ideas will be implemented, and with what speed. When there are many individuals systematically engaged in the search for new ideas the likelihood is increased that an invention will be readily adopted. Thus, there is positive feedback between the number of inventors and innovators, with the possibility of a high-level equilibrium and a low-level equilibrium.[15] Institutionally, this is partly reflected in the existence of organizations and firms, such as venture capital firms in more developed countries, who specialize in searching for and financing new ideas.

Thus, a greater 'supply' of innovation leads to the development of institutions which facilitate innovation. But there are other ways by which an increase in the amount of innovation increases an economy's innovative capacity. The ability to learn is itself learned; though some learning occurs in educational institutions, most learning occurs in social contexts, in which no charge is imposed for the learning externality.

There may, of course, also be a negative feedback: with a fixed stock of ideas, the greater the number of individuals engaged, say, in a patent race, the lower the expected return to each. Our hypothesis is that, at least in many situations, the positive feedbacks dominate this negative feedback.

Finally, the nature of any organization depends on the demands imposed on that organization for adaptation, which is itself a function of the society's state of innovativeness.[16] The supply of innovation gives rise, in a sense, to its own demand. In unchanging environments, there is little demand for individuals and institutions who know how to cope with changes, those who specialize in the collection, processing and dissemination of information. Further, the return to searching for new ideas may be lower, simply because there are fewer ideas floating around; there is a certain synergy in the production of new ideas.

Note that the model outlined in the previous section captures those types of interactions where market or public information is relatively scarce or has low accuracy.[17] On the other hand, many of the kinds of positive feedbacks we have discussed in this section would exist even if there were publicly available accurate information, say about the number and details of venture capital firms, and the variables representing the demands and remunerations associated with various kinds of jobs and qualifications. The information structure of the economy (what information is collected and how it is transmitted) obviously affects the stochastic distribution of information, and this, in turn, may affect not only the dynamics, but even the nature of the long-run steady-states to which the economy converges.

Still, in all of these models, the effect of the past is felt, though the nature and the magnitude of this effect will be determined by the nature of the available information-generating processes. The positive feedbacks which we discussed in the preceding paragraphs still give rise to multiple steady-states; there may be no long-run convergence either in levels of income or growth rates.

3.1 An Example

The following simple model provides an illustration of this non-convergence. As we noted in section 1 it is easier to construct multiple equilibria assuming differences among individuals; here we assume that for some individuals the cost of being innovative is greater than for others. As in the model of section 2 we shall focus on situations where the returns to pursuing a particular action depend on the actions taken by the individuals with whom one interacts; we shall also assume that there is a sufficiently large number of individuals so that no one believes that his actions will have any effect on the actions taken by others, that is, we ignore all strategic considerations. Unlike section 2, we shall (at least initially) assume individuals know the fraction of individuals undertaking different actions.

For simplicity we assume that there are two strategies, $\propto$ (innovative) and β (non-innovative or bureaucratic). Individuals interact pair-wise. When individual i pursues strategy $\propto$ and individual j, with whom he interacts, pursues strategy β, the pay-off to individual i is $\pi(\propto,\beta)$. When both pursue strategy $\propto$, the pay-off is $\pi(\propto,\propto)$; when both pursue β it is $\pi(\beta,\beta)$. We assume that while individuals have the same 'gross' pay-off function, individuals differ in the cost of undertaking strategy $\propto$ versus strategy β. Assume that pursuing $\propto$ costs c dollars more than pursuing β, and that the distribution of c in the population is given by $F(c)$. Assume that individuals believe that a fraction θ of the population is pursuing strategy $\propto$. Then the expected gross pay-off to pursuing strategy $\propto$ is

$$\theta\pi(\propto,\propto) + (1-\theta)\pi(\propto,\beta)$$

All of those for whom $c < c^*$ will pursue $\propto$, where c^* is defined by

$$c^* = [\theta\pi(\propto,\propto) + (1-\theta)\pi(\propto,\beta)] - [\theta\pi(\beta,\propto) + (1-\theta)\pi(\beta,\beta)] \tag{8}$$

Then equilibrium is defined by

$$c^* = \{F(c^*)\pi(\propto,\propto) + [1-F(c^*)]\pi(\propto,\beta)\} - \{F(c^*)\pi(\beta,\propto) + [1-F(c^*)]\pi(\beta,\beta)\} \tag{9}$$

or, alternatively,

$$\theta^* = F[c(\theta^*)] \tag{10}$$

where $c(\theta)$ is the value of c associated with that individual who is indifferent between the two strategies when a proportion θ of the population pursue strategy $\propto$. At θ^*, the number of individuals pursuing strategy $\propto$ makes it worthwhile for precisely that fraction, θ^*, to pursue strategy $\propto$. A corner equilibrium is defined by

$$\pi(\propto,\beta) - \pi(\beta,\beta) < c_{\min}$$

When everyone is pursuing β, it pays the individual for whom the cost of pursuing $\propto$ is the least to pursue β. Another equilibrium is defined by

$$\pi(\propto,\propto) - \pi(\beta,\propto) > c_{\max}$$

When everyone is pursuing $\propto$, it pays the individual for whom the cost of pursuing $\propto$ is the highest to pursue $\propto$.

It is apparent that if

$$\pi(\propto,\propto) - \pi(\propto,\beta) - \pi(\beta,\propto) - \pi(\beta,\propto) + \pi(\beta,\beta) > 0$$

then as more individuals pursue $\propto$, the expected return to pursuing $\propto$ increases. There may exist multiple equilibria as illustrated in figures 19.4 and 19.5. A sufficient condition for the existence of multiple equilibria is that

$$\pi(\propto,\beta) - \pi(\beta,\beta) < c_{\min} \tag{a}$$

$$\pi(\propto,\propto) - (\propto,\beta) > c_{\max} \tag{b}$$

(a) and (b) ensure that there exist two corner equilibria.

The analysis so far has focussed on steady-states. If for generations the

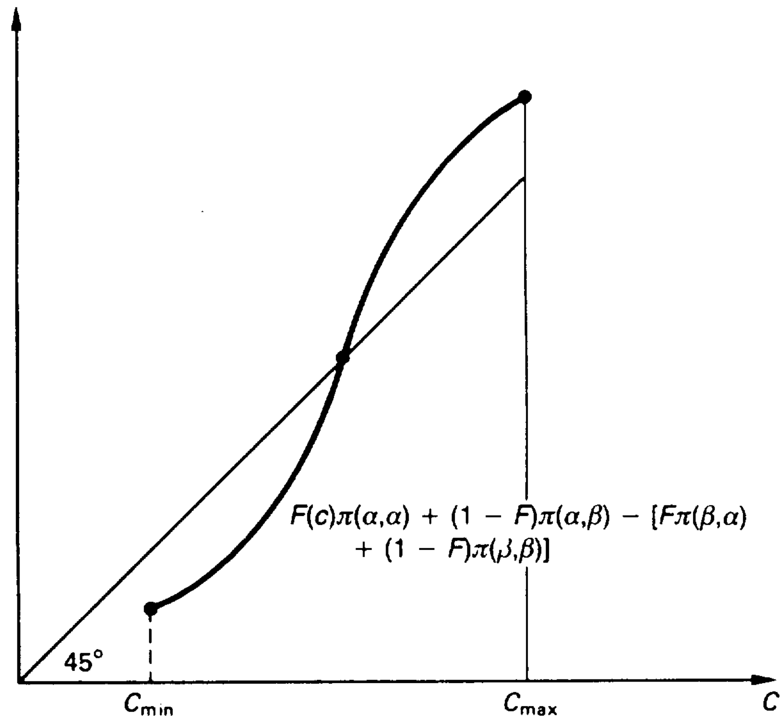

Figure 19.4 Multiple equilibria in willingness to innovate (corner solutions)

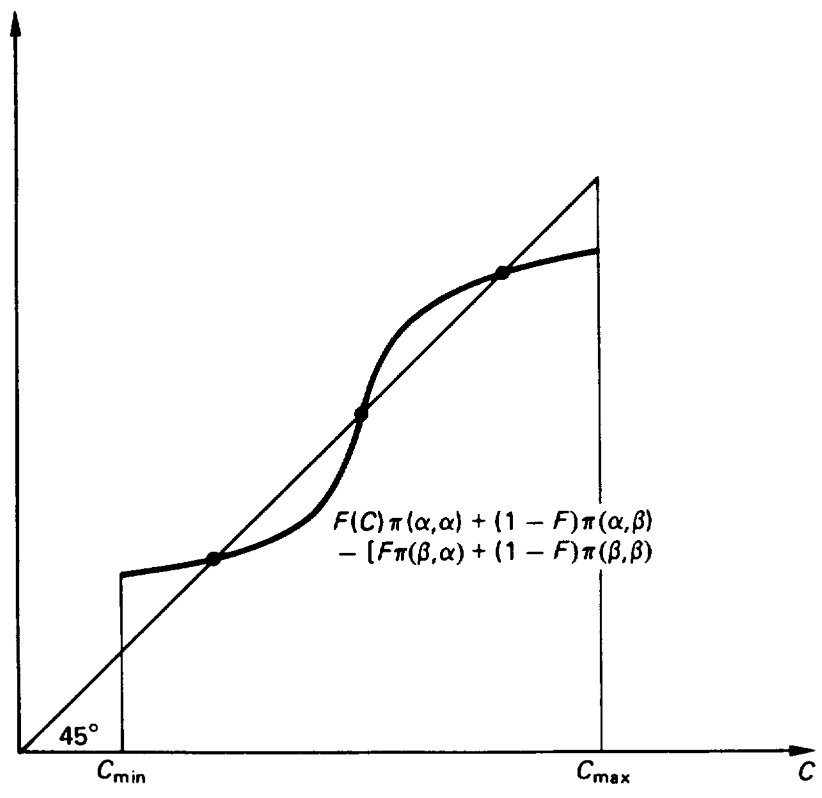

Figure 19.5 Multiple equilibria in willingness to innovate (interior solutions) fraction of those who have acted innovatively is θ^*, it is plausible that individuals both know this and come to expect that the chances that the next individual they encounter will act innovatively is θ^*. But this does not explain how the equilibrium is attained. How do individuals come to know $\theta(t)$, and, even were it possible for them to know $\theta(t)$, how do they use that information to predict what $\theta(t+1)$ will be?

The great strength of the models discussed in section 2 is that they explicitly model a learning process. The particular learning process modeled has the property that, even in the steady-state, individuals never learn about the aggregate value of θ^* defined above. In some contexts, however, it is plausible to assume that the dispersion of beliefs about the aggregate value of θ decreases over time, and that a model similar to the one described in this section may provide a good description of the steady-states.

4. Other Perspectives

We have repeatedly shown externalities associated with technical change giving rise to positive feedbacks, which in turn give rise to multiple innovation/bureaucratic equilibria. Externalities similar to those discussed earlier can arise from several other sources. In this section, we analyze situations in which the distribution of characteristics in the population is affected not by decisions, but by either processes of taste acquisition or natural selection.

4.1 Tastes

It is not only capacities and beliefs which are learned, as was emphasized in the previous two sections, but also tastes. Tastes for innovation or for routinization can be acquired. Scitovsky (1976) has emphasized the basic psychological tension between 'novelty' and 'excitement' on the one hand, and 'comfort' or 'routineness' on the other. Those who like change become impatient in an environment in which they are asked to do repetitive, routine tasks. On the other hand, workers who are accustomed to routine tasks often complain when they are assigned to jobs requiring non-routine behavior. How the tension between novelty and routine gets resolved can be affected by the environment. A basically innovative personality placed in a bureaucratic environment may become a routine lover.

Though the details of the mechanisms of preference formation and change are not fully understood, many observers have noted that intrinsic preferences change,[18] particularly in intra-organizational contexts. Consider, for instance, an intra-organizational context where there are two types of managers, innovators and bureaucrats. While the latter enjoy routines and well-established patterns of behavior, the former enjoy change. If tastes are acquired then junior workers will tend to duplicate their superiors' preferences. Different kinds of preferences can, therefore, generate externalities.

4.2 Birth and Death Processes

Another aspect of organizations which we now examine is that new organizations appear to inherit some of the traits of the organizations which have spawned them. That is, an organization spawned by an innovative organization will be more innovative than an organization spawned by a bureaucratic organization. Note, however, that since the creation of a new organization is, in some sense, an innovative act, it seems reasonable that new organizations are more likely to be innovative than bureaucratic, and as organizations age they are more likely to become bureaucratic.

Let x denote the fraction of innovative organizations. Let $b^1(x)$ and $b^2(x)$ denote the birth rates innovative firms associated respectively with an innovative and a bureaucratic firm. Let $d(x)$ denote the death rate of innovative firms. Then equilibrium in the fraction of innovative firms is represented by

$$xb^1(x) + (1-x)b^2(x) = xd(x) \tag{11}$$

which can be restated as

$$x = b^2(x)/[d(x) - b^1(x) + b^2(x)] = c(x) \tag{12}$$

The derivatives of d, b^1 and b^2 (denoted respectively by d_x, b^1_x and b^2_x) have ambiguous signs in general. For instance, if innovative firms simply become bureaucratic with age, at some constant rate, then $d_x = 0$. If the number of innovative firms increases the effective degree of competition, then $d_x > 0$. If, on the other hand, bureaucrats make the life of innovators more difficult, $d_x < 0$. The sign of b^2_x is also ambiguous. While it is reasonable to postulate that an increase in the fraction of innovative firms decreases the profitability of bureaucratic behavior, it is not apparent whether low profits of bureaucratic firms spawn more births of innovative firms (because the relative gains to those with more innovative characteristics striking out on their own becomes more apparent), or whether low profits spawn fewer births of innovative firms (because the apparent lack of profitable opportunities attracts less entry).

In general, therefore, $c(x)$ can take on a variety of shapes. It can be, at least through a range, upward sloping (similar to the curve in figure 19.1), giving rise to multiple social equilibria. There may be a stable bureaucratic equilibrium, an equilibrium with a preponderance of bureaucratic firms and another stable innovative equilibrium with a preponderance of innovative firms. This is because, within the bureaucratic environment, it simply does not pay to be very innovative; and within innovative environments, bureaucrats cannot easily survive.

5. The Technological Perspective

The simple neoclassical model, which predicts a convergence of rates of growth, employs a number of stringent assumptions. We do not wish to overemphasize the contrast between the technological perspective and the perspectives discussed in previous sections. In particular, several of the arguments for positive feedbacks were based on technological considerations, on the externalities which arise between the amount of innovation in one sector and productivity in another. Our focus, here, is on certain properties of economic aggregates, rather than the behavior of individual units (individuals, organizations) which was the center of concern in preceding sections. Research in this area has attempted to identify what alternative technological assumptions give rise to multiple equilibria, e.g., one equilibrium in which there is a high rate of technological change and another in which there is a low rate.

Three aspects of technology which give rise to multiple equilibria have been identified.

1 Much of technological learning is a by-product of production (Arrow, 1962). Learning curves associated with different products and different technologies may differ markedly.
2 Much of technological learning is localized. That is, accretions in knowledge that are relevant to one technology may have little bearing on other technologies. Spill-overs are far from uniform (Atkinson and Stiglitz, 1969).
3 The process of learning is, itself, learned. Thus, just as Adam Smith emphasized the importance of specialization in production, the importance of specialization in learning has become increasingly apparent. Just as the worker who specializes in producing pins becomes more proficient in pin production unless he suffers from boredom, so too might the individual who specializes in research become more proficient in doing that.[19]

These observations concerning technological learning have important implications for economic theory and policy.[20] Among them are the following.

5.1 Multiplicity of Equilibria

It is possible to show that one consequence of the above technological hypotheses is that there may be multiple long-run equilibria. Assume, for instance, that more capital-intensive technologies have a greater capacity for learning.[21] That is, the rate of increase in productivity associated with any increment in production is greater in such technologies. Then some economy may be trapped in a low-level equilibrium, with a low capital–labor ratio and a low rate of technological progress, even though there exists another equilibrium with a high capital–labor ratio and a high rate of technological progress. Thus, this model provides an explanation of the non-convergence of rates of growth as well as levels of income.

There is one problem with this explanation. In the context of an international economy in which knowledge, and individuals who have the knowledge, can move across national borders, less capital-intensive economies have potential access to the information available in more developed countries. At the same time, however, the theory of localized technological change may provide a part of the explanation for why such knowledge, developed by more developed economies, may have limited relevance for LDCs. The latter face the direct costs of adapting these technologies to local conditions, such as a higher variance in the quality of inputs. Those with the less capital-intensive technologies may be deterred from using technologies developed in the more developed countries.

Learning-to-learn effects strengthen the latter set of conclusions. Previous lack of experience may limit the capacity of individuals in some economies to take full advantage of possibilities for productivity improvement associated with certain technological changes.

Finally, if changes in technology in the more developed countries make it increasingly difficult for those in LDCs to benefit (even with a lag) from the advances, either because they require technical skills on the part of workers which are scarce or because they require a degree of quality control in inputs which LDCs find hard to attain, then there may be a growing gap in living standards and productivity between the poorest of the LDCs and the more developed countries.[22]

5.2 Non-convexities and the Disadvantages of Being Late

As Arrow recognized in his early essay, learning-by-doing gives rise to non-convexities. These non-convexities are exacerbated by the presence of localized learning and learning-to-learn effects.

Non-convexities are important for at least two reasons. First, they imply that specialization is often advantageous. While, in the absence of non-convexities, a country might pursue a gradual process of transition from, say, labor-intensive to capital-intensive technologies, non-convexities may make this gradual approach less desirable, providing an alternative basis for a 'Big-Push' theory of development. The fact that a gradual transition is not desirable may serve as an impediment to development for the very poorest countries: in order to undertake a Big-Push, greater sacrifices in current consumption are required than these countries might be willing to make.[23]

Moreover, the sunk costs invariably associated with technological progress, whether arising from learning effects or expenditures on R&D, create an entry barrier which puts latecomers (the LDCs) at a marked disadvantage. In the presence of sunk costs, potential competition will not in general drive profits (rents on existing innovations) to zero (Stiglitz, 1988); it is certainly possible that there will be periods in which innovation rents increase, and accordingly, in which the gap in income between the innovating countries and the non-innovators may increase.

6. Concluding Remarks

6.1 A General Mathematical Formulation

There is a simple mathematical model underlying all of the formulations discussed in sections 1 to 5. Consider an economy in which individuals are characterized by some variable (say the degree of innovativeness). For simplicity, we shall assume that the variable can take on only discrete values, and for concreteness, we shall further assume it can take on only two values. The proportion of the population with characteristic i at date t is given by $x_i(t)$. Clearly, $0 \leq x_i(t) \leq 1$ and $\Sigma x_i = 1$. Thus the economy at date t is described by the vector $x(t) = \{x_1(t), x_2(t), \ldots\}$ representing the distribution of characteristics.

The economy is characterized by certain transition rules, which specify the proportion of those with characteristic i at date t who have (or whose descendants have) characteristic j at time $t+1$. These transition rules themselves may be – and in general will be – a function of the population mix at time t (and possibly at previous periods as well). Thus, we write,

$$x(t+1) = A[x(t), x(t-1), x(t-2), \ldots]x(t)$$

If A were independent of x then, under standard conditions, there would exist a unique steady-state equilibrium vector x^* satisfying

$$x^* = Ax^*$$

But the environment – here, the vectors $\{x(t), x(t,-1), \ldots\}$ – determines the nature of the transitions. A generally depends on the x's, and accordingly, there may be multiple solutions to the equation

$$x^* = A(x^*, x^*, \ldots)x^*$$

each of which can be thought of as a social equilibrium.[24]

The above description is obviously too general. We need to derive the transition rule from structural models, to explain why, and the circumstances in which, the proportion of individuals in the population with a given characteristic increases or decreases. That is precisely what the various models discussed in previous sections attempt to do.

6.2 Strategic Considerations

It should be apparent that the theories of social equilibrium we have put forth differ in a fundamental conceptual way from standard game theory, in which multiple equilibria may arise as well. In typical game theoretic models, individuals arrive at beliefs about what their rivals will do, and therefore about what is optimal for them to do, on the basis of introspection. They know the pay-offs, and the set of available strategies, and they believe that their opponents are rational. They can infer, therefore, what their rivals will do. However, some critical problems remain. For instance, there are no satisfactory theories describing the circumstances which lead to different possible Nash equilibria, or indeed, how equilibrium is attained in the presence of multiple Nash equilibria.

We would argue that most real-life situations are too complicated for individuals to feel confident, simply through introspection, about what their rivals will do. They seldom know their rivals' pay-offs, and they may not even be sure about the set of strategies available to those with whom they interact. This is particularly true in contexts where innovation is important, where actions entail thinking up something that has not been done before. But even when the pay-offs and available strategies are known ('common knowledge'), it may be difficult to infer what one's rival will do: even a simple game like chess is too complicated to 'solve.' Moreover, recent arguments by Reny (1988) and Binmore (1985) cast doubt on the underlying hypotheses of standard game theory. They have identified a variety of important situations in which the postulates of rationality and common knowledge are inconsistent.

6.3 Externalities and Multiplicity

There is an old saying that some of the best things in life are free. Only a fraction of our interactions are completely regulated by the price system. Any parent recognizes that his child learns attitudes not only from his parents but from peers, and from a wide variety of environmental influences. More than that, one picks up modes of thought and behavior. In all cultures, information is constantly exchanged among individuals; but the nature of the information exchanged may differ. In Silicon Valley in California, information about the most recent developments in computers is exchanged; on Main Street, on the other hand, it may be the latest piece of juicy gossip. Whether economically motivated or not, such interactions may have a profound effect on economic behavior and equilibrium. There are, to use the economist's traditional jargon, important externalities arising out of these social interactions. This is not to say that social interactions are not affected by economic returns: the direction of conversation – the greater talk about computers in Silicon Valley – may reflect the greater economic returns there to such information. Still, the central point that we have emphasized, the possibility of multiple equilibria, remains.

What is required for success in the development process is the transformation of a society into an innovative and adaptive culture; and what is required for that is more than the shipment of capital, the construction of oil refineries and tire factories. Indeed, some of the central roles that the government has played (for instance, in creating a large, entrenched and privileged bureaucracy attracting a significant part of the economy's pool of talent) may have served, in the long run, to suppress the development of such a culture. If this argument is correct, it has profound implications for how we should think of development strategies.

NOTES

Support from the National Science Foundation and the Hoover Institution is gratefully acknowledged.

1 Even if LDCs adopt the best practices of developed countries with a lag, according to this view, the *rates* of technological progress will be the same and differences in levels of per capita income will then be related also to the length of the lag in the diffusion of technology.
2 Both authors have had the good fortune of interacting with Carlos Díaz-Alejandro. Stiglitz's interests in development were stimulated by him, first as colleagues at Yale and later through the contact they kept up. Sah had the pleasure of discussing with Carlos Díaz-Alejandro some of the mysteries of economic history. Our emphasis on his intellectual contributions is not meant to understate his many other rich dimensions. He brought to all aspects of his life the same liveliness, humor, excitement, openness, breadth of perspective and sense of balance combined with a seriousness of purpose and a sense of commitment that he brought to his academic work.
3 The difficulty of ruling out the possibility of such multiple equilibria becomes much greater when many factor and commodity markets are considered together.
4 Though these theories did not articulate precisely how they differed from the standard Arrow–Debreu model (nor did they see the need to do so, since the point they made was perfectly clear), the central importance of the assumption of an incomplete set of markets should be clear. The users of steel would not find it profitable to construct their plants to make steel-using products in the absence of the availability of steel, while the producers of steel would not find it profitable to manufacture steel in the absence of users. Stiglitz (1987) discusses the conditions which make it likely that this kind of externality can be internalized.
5 We require $\gamma_1 - w(0) < c < \gamma_1 - w(1)$, a condition which can easily be satisfied. In these circumstances, there exists an interior equilibrium which, under most natural specifications of dynamics, is unstable.
6 Again, we can Pareto rank the equilibria. The difference between the incomes of the more able certified and uncertified is fixed; and as more individuals are certified, the wage of the uncertified decreases so that equilibria with fewer certified individuals will Pareto dominate.
7 Multiplicity of equilibria obviously arise in many other contexts as well. Diamond (1982) and Drazen (1987) have emphasized the importance of positive feedback in search models. If more employers are looking for employees, it pays workers to search more; and if more workers are searching, it may pay more employers to search.
8 That will be the case if θ is sufficiently small. If m is small, or if γ_1 is sufficiently small, then it is possible that this condition will be satisfied for all values of θ.
9 The latter is particularly important, given the possibility, noted already, of Pareto-inferior equilibria.
10 An extensive literature deals with expectation-formation processes. For alterna-

tive views on rational expectations, for instance, see Frydman and Phelps (1983).

11 Slight modifications of the model deal with cases in which the individual makes many choices each period, or in which longer term choices are entailed because, for instance, there are fixed costs associated with choices.

12 These utilities obviously vary across individuals. Also, they depend on economy-wide variables; for example, U's depend on $y(t)$. The conclusions drawn below hold if, within the relevant ranges of the parameters and for most (but not all) individuals, the reward function exhibits the property assumed in the text.

13 Some of these steady-states may be stable while others may be unstable. Sah shows that some of the sufficient conditions which guarantee stable interior steady-states are extremely mild. Explicit relationships such as (7) also provide a basis for analyzing comparative statics of steady-states with respect to changes in parameters. For previously unavailable results on such comparative statics of systems of dynamic equations, see Sah (1987b).

14 The number of individuals who are innovative acts like an atmospheric externality.

15 It is, of course, possible that diminishing returns set in. With a fixed stock of ideas, the marginal contribution of an additional innovator may diminish the gain of a number of other innovators. We are suggesting that, in many instances, this effect may be outweighed by the externality effect which is the unappropriated contribution of any innovator to the stock of available ideas.

16 This is illustrated by recent experiences in the US airline industry. In the days of airline regulation, some airlines developed organizational structures suited for dealing with the Federal bureaucracy; these thrived, and there was little room for more market-oriented airlines. Following deregulation, the latter thrived, while the former, not well suited to the competitive environment, languished.

17 Other aspects of such models, including their similarities and differences with models which have been used in various other contexts, are discussed in Sah (1988b).

18 That is, these changes are distinct from, for instance, changes in long-term choices which individuals make as a result of their beliefs or pecuniary trade-offs.

19 There are obvious dangers from excessive specialization. Specialized individuals may be less able to adapt to changes that are sufficiently far removed from their specialization, though their greater specialized knowledge may make them more able to adapt to changes that are within their sphere of specialization. The trade-offs between specialization and adaptability have not yet been adequately studied.

20 Related concerns are discussed in Lucas (1985), Romer (1986) and Sah and Stiglitz (1988).

21 Lucas has emphasized the differences in learning capacities associated with different products. He assumes limited spill-overs across national boundaries, but perfect spill-overs within a country [otherwise, as Dasgupta and Stiglitz (1988) show, there cannot exist a competitive equilibrium]. Accordingly, equilibrium will be associated with different countries specializing in different commodities. The more developed countries have the advantage of being able to choose those commodities with better learning curves. But what is important is not rates of growth in physical productivity, but in incomes; changes in relative prices may partially, or more than, offset these differences. Thus, if all individuals had unitary price elasticities for all commodities, changes in relative prices would precisely offset changes in relative productivities. See Skeath (1988).

22 To put it another way, the lag between the state of technology of the developed and the less developed countries is *endogenous*, and changes in technology have increased that lag.

23 That is, not only may multiple steady-states arise in a *descriptive* growth model, but there may be multiple steady-states in an optimal growth model. To which steady-state one converges depends on the initial conditions.

24 If the population consists of individuals of various ages, then the dynamics can be described by the fraction of those who are of age t and type i who become type j at age $t + 1$. There is thus a transition matrix for each age cohort. This approach provides a much richer description of the dynamics of the economy than an approach which simply focusses on the aggregates. See for example Sah (1987a).

REFERENCES

Arrow, K. J. 1962: The economic implications of learning by doing. *Review of Economic Studies*, 29, 155–93.

Atkinson, A. B. and Stiglitz, J. E. 1969: A new view of technological change. *Economic Journal*, 79, 573–8.

Baumol, W. J. 1986: Productivity growth, convergence and welfare: what the long-run data show. *American Economic Review*, 76, 1072–85.

Binmore, K. G. 1985: Modeling Rational Players. Mimeo, London School of Economics and University of Pennsylvania.

Dasgupta, P. and Stiglitz, J. E. 1985: Sunk Costs, Competition and Welfare. Mimeo, Princeton University.

Dasgupta, P. and Stiglitz, J. E. 1988: Learning-by-doing, market structure, and industrial and trade policies. *Oxford Economic Papers* (forthcoming).

Debreu, G. 1959: *Theory of Value*. New York: Wiley.

Diamond, P.A. 1982: Aggregate demand management in search equilibrium. *Journal of Political Economy*, 90, 881–94.

Drazen, A. 1987: Reciprocal externality models of low employment, *European Economic Review*, 31, 436–43.

Frydman, R. and Phelps, E. S. (eds.), 1983: *Individual Forecasting and Aggregate Outcomes*. Cambridge: Cambridge University Press.

Lucas, R. E. Jr. 1985: On the Mechanics of Economic Development, Marshall Lectures, Cambridge University.

Reny, P. J. 1988: Common knowledge and extensive form games. *Journal of Economic Perspectives* (forthcoming).

Romer, Paul 1986: Increasing returns and long run growth. *Journal of Political Economy*, 94, 1002–38.

Rosenstein-Rodan, P.N. 1943: Problems of industrialization of Eastern and South-Eastern Europe. *Economic Journal*, 53, 204–7.

Sah, R. K. 1985: What Affects the Level of Honesty in an Economy. Working Paper Series D, School of Organization and Management, Yale University, New Haven.

Sah, R. K. 1987a: Persistence and Pervasiveness of Corruption: New Perspectives. Paper presented at the Conference on Political Economy, The World Bank, Washington, DC, June 1987.

Sah, R. K. 1987b: Results for Economic Comparative Statics of Steady-States of Higher-Order Discrete Dynamic Systems. Working Paper, Yale University, New Haven.

Sah, R. K. 1988a: Social Osmosis and Patterns of Crime: A Dynamic Economic Analysis. Working paper, Yale University, New Haven.

Sah, R. K. 1988b: Social Osmosis and Patterns of Observed Economic Behavior: Crime, Corruption and Some Other Social Phenomena. Working paper, Yale University, New Haven.

Sah, R. K. and Stiglitz, J. E. 1988: Technological Learning, Social Learning and Technological Change. Paper presented to the World Congress of the International Economic Association, New Delhi, December 1986. Forthcoming in proceedings.

Scitovsky, T. 1976: *The Joyless Economy*, New York: Oxford University Press.

Skeath, S.E. 1988: Learning, Price Effects and Growth. Mimeo, Princeton University.

Solow, R. M. 1956: A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70, 65–94.

Starrett, D. 1976: Social institutions, imperfect information and the distribution of income. *Quarterly Journal of Economics*, 90, 216–84.

Stiglitz, J. E. 1974: Theories of discrimination and economic policy. In G. M. von Furstenberg et al. (eds.), *Patterns of Racial Discrimination*, vol. II. Lexington: Lexington Books.

Stiglitz, J. E. 1975: The theory of 'screening' education, and the distribution of income. *American Economic Review*, 283–300.

Stiglitz, J. E. 1984: Information, screening and welfare. In M. Boyer and R. Khilstrom (eds.), *Bayesian Models in Economic Theory*, Amsterdam: North–Holland, pp. 209–39.

Stiglitz, J. E. 1987: Learning to learn, localized learning and technological progress. In P. Dasgupta and P. Stoneman (eds.), *Economic Policy and Technological Performance*, Cambridge: Centre for Economic Policy Research, Cambridge University Press, pp. 125–53.

Stiglitz, J. E. 1988: Technological change, sunk costs and competition. *Brookings Papers on Economic Activity* (forthcoming).