ENDOGENOUS BUSINESS CYCLES: SOME THEORY AND EVIDENCE

by

MOHAMAD LUTFI HAMMOUR

B.A., Philosophy
Stanford University
(1985)

M.S., Industrial Engineering
Stanford University
(1985)

Submitted to the Department of Economics
in Partial Fulfillment of
the Requirements of the Degree of
Doctor of Philosophy

at the

Massachusetts Institute of Technology

September 1989

Signature of Author__________________________________
Department of Economics
July, 1989

Certified by__________________________________
Olivier J. Blanchard
Professor of Economics
Thesis Supervisor

Accepted by__________________________________
Professor Peter Temin, Chairman
Departmental Graduate Committee
Department of Economics

ENDOGENOUS BUSINESS CYCLES: SOME THEORY AND MEASURMENT

by

Mohamad Lutfi Hammour

Submitted to the Department of Economics on July 10, 1989
in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

ABSTRACT

The three essays that make up this thesis are a contribution toward an *endogenous* theory of economic fluctuations. The first two essays advance some theoretical arguments why the presence of "aggregate increasing returns" can cause macroeconomic instability and place the economy in a state of self-sustained fluctuations. The third essay is an attempt to detect such instability in post-war U.S. data using nonlinear times series estimation techniques.

Essay 1 argues in the context of the basic neoclassical models of savings and investment that economic dynamics can become essentially endogenous in the presence of increasing returns. Under constant returns to scale, the economy converges to a saddle path stable equilibrium. But under either internal or external economies of scale, increasing returns destabilize the equilibrium. Moreover, if there are natural "expansion diseconomies," the economy is likely to converge to a stable limit cycle around steady state. Together with endogenous business cycles, increasing returns may also result in multiple equilibria of two sorts: indeterminacy, with a possible role for "animal spirits"; and multiple steady states, with the related "coordination failure" problems.

Essay 2 generalizes these ideas to an abstract class of general equilibrium models with external effects of aggregates on individual decisions, that includes recent macro models of imperfect competition, search, etc. These externalities can potentially generate multiple equilibria, with the resulting coordination failure problems and a role for "animal spirits." Cooper and John have shown the necessity of "strategic complementarities" for multiplicity to arise. This essay derives the stronger necessary condition of "aggregate increasing returns," and discusses the way it arises in specific examples ("thick-market" externalities, increasing returns in production, countercyclical markups). It explores this condition's dynamic implications, and shows that it has an essentially destabilizing effect. Different limits on expansion are discussed that can keep the economy fluctuating endogenously within the unstable region.

Essay 3 turns to the data, and attempts to discriminate empirically between endogenous and exogenous theories of the cycle. It takes advantage of the well known fact that the "exogenous shocks" approach to business cycles implicitly assumes the economy tends to converge to a stable steady state, while the endogenous approach assumes its steady state is unstable. A simple nonlinear time series model is used to estimate the local dynamics around the economy's steady state and discriminate between the two approaches. Using the post-war U.S. industrial production index, the hypothesis of steady-state stability can be rejected: It is found that the economy has tended to diverge away from its steady-state growth path in periods when it was close to it. A similar conclusion is reached using employment data.

Thesis Supervisor: Olivier J. Blanchard
Title: Professor of Economics

بسم الله الرحمن الرحيم

ACKNOWLEDGMENTS

I am grateful to my advisors, Olivier Blanchard and Robert Solow, for their help, encouragement, and advice. I also thank Jordi Gali, Peter Howitt, Sung-In Jun, Winston Lin, Kevin Murphy, Danny Quah, Julio Rotemberg, Robert Solow, Martin Weitzman, Michael Woodford, and Jeff Woolridge for useful discussion and comments regarding various parts of this thesis. Responsibility for all remaining errors is solely mine.

INTRODUCTION

The three essays that make up this thesis are a contribution toward an *endogenous* theory of economic fluctuations. From a theoretical point of view, the goal is to propose an *expalnation* for observed macro instability—if any such explanations are relevant at all—that goes beyond postulating the presence of "exogenous shocks." Empirically, one would like to find ways to discriminate between exogenous and endogenous theories of the cycle.The first two essays advance some theoretical arguments in favor of the possibility of self-sustained fluctuations, based on the concept of "aggregate increasing returns." The distinguishing feature of this line of reasoning is that, unlike some previous endogenous business cycle models, it is fully grounded on the neoclassical foundations of individual optimization in a general equilibrium context. The third essay is a preliminary attempt to discriminate empirically between exogenous and endogenous theories by testing for the steady-state stability of the dynamic system governing observed aggregate time series.

Essay 1 argues in the context of the basic neoclassical models of savings and investment that economic dynamics can become essentially endogenous in the presence of increasing returns. Under constant returns to scale, the economy converges to a saddle path stable equilibrium. But under either internal or external economies of scale, increasing returns destabilize the equilibrium. Moreover, if there are natural "expansion diseconomies," the economy is likely to converge to a stable limit cycle around steady state. Moreover, it is shown that, together with endogenous business cycles, increasing returns may result in multiple equilibria of two sorts: indeterminacy, with a possible role for "animal spirits"; and multiple steady states, with the related "coordination failure" problems.

Essay 2 is a generalization of these ideas. It focuses on the important role of external effects of aggregates on individual decisions in generating multiple equilibria and self-sustained fluctuations. It takes as a starting point Cooper and John's (1988) result concerning the necessity of "strategic complementarities" for multiplicity to arise in macro models with external effects. It then derives the stronger necessary condition of "aggregate increasing returns," and discusses the way this arises in specific examples. It explores this condition's dynamic implications, and shows that it has an essentially destabilizing effect on the economy. Different limits on expansion are discussed that can keep the economy fluctuating endogenously within the unstable region.

The abstract nature of these results is such that they can appear in many economic examples. Aggregate increasing returns can be due to "thick-market" externalities, increasing returns in production, or countercyclical markups. They can affect such diverse individual decisions as capital investment, inventory accumulation, or hiring. Only a detailed empirical investigation could reveal if any of these variables is subject to aggregate increasing returns, in which case it promises to provide a key to explaining macro fluctuations.

Finally, essay 3 turns to an empirical attempt to discriminate between endogenous and exogenous theories of the cycle. It takes advantage of the well known fact that the "exogenous shocks" approach to business cycles implicitly assumes the economy tends to converge to a stable steady state, while the endogenous approach assumes its steady state is unstable. A simple nonlinear time series model is used to estimate the local dynamics around the economy's steady state and discriminate between the two approaches. Using the post-war U.S. industrial production index, the hypothesis of steady-state stability can be rejected: It is found that the economy has tended to diverge away from its steady-state growth path in periods when it was close to it. A similar conclusion is reached using employment data. However promising they may be, these results are very preliminary. Much more work needs to be done using richer nonlinear times series models and additional data before any attempt at a conclusion is made.

The arguments in this thesis provide promising clues for explaining observed macro instability, but only further empirical and theoretical investigation can tell whether these clues will turn out to be useful.

REFERENCE

COOPER, R., AND A. JOHN (1988): "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics*, 103, 441-64.

ESSAY 1

INCREASING RETURNS
AND
ENDOGENOUS BUSINESS CYCLES

One way to classify theories of the business cycle is to distinguish between endogenous and exogenous theories[1]. Endogenous theories hold that the cycle is the result of a deep *structural instability* in the economy. Certainly external shocks tend to exacerbate the situation, but fluctuations would still take place in their absence. Nonlinear multiplier-accelerator models are notable representatives of this approach (Hicks (1950), Goodwin (1951)). Exogenous theories, on the other hand, give much more importance to exogenous shocks. They assert that the economy tends naturally to converge to a *stable* equilibrium and that the only way it can exhibit the observed sustained fluctuations is for it to be constantly hit by exogenous shocks that keep it away from steady state. This approach was formulated in the early work of Slutzky (1937) and Frisch (1933).

In principle one may object to the relevance of this distinction since what is considered exogenous is a matter of the extent of the model. Two models may consider the same factor to be the cause of the cycle, but one may include it among the factors it is explaining and take it as endogenous, while the other may not and thus take it as exogenous. In practice, however, most models take as endogenous those factors within the private economy that have an economic explanation—mainly, the behavior of households and firms—and as exogenous those that are not—such as wars, technological breakthroughs, government policy—or that originate outside the economy, such as the oil shocks of the 1970s. The real question is, therefore, which of these factors are behind the observed instability in the economy.

In the more recent literature the Slutzky-Frisch framework has become the paradigm for thinking about the cycle while endogenous theories have gone out of fashion. The general adoption of modern linear time series analysis as the fundamental technique for macroeconometric analysis greatly contributed to this state of affairs because, by its very assumption of linearity, it almost entirely precludes the presence of autonomous fluctuations. From a theoretical point of view, the implicit assumption of steady-state stability underlying this approach is most often justified by the fact that it holds in the "basic neoclassical model" of capital accumulation (i.e., the decentralized version of the Ramsey (1928) model) and its derivatives[2].

In this essay, I argue that this stability result hinges critically on the neoclassical assumption of constant returns to scale. If this assumption is relaxed and we allow for *increasing returns*, the economy's steady state easily loses its stability and a business cycle can arise endogenously. Thus, once we recognize the widespread existence of increasing

---

[1]A good survey of the literature can be found in Zarnowitz (1985). See also Gabisch and Lorenz (1987).
[2]See Kydland and Prescott (1982).

returns in industrial economies (see, e.g., Pratten (1971), Hall (1988)), we must take the possibility of endogenous fluctuations seriously and can no longer limit ourselves to an exogenous view of the cycle.

The idea that economic dynamics become essentially endogenous in the presence of increasing returns can be traced back at least to Allyn Young (1928) who formulated it in the context of growth theory, and was later taken up by Nicholas Kaldor (1972) who put it as follows:

> Abandoning the axiom of "linearity" and assuming that, in general, the production of any one commodity, or any one group of commodities, is subject to increasing returns to scale [implies that we can no longer maintain that] economic forces operate in an environment that is "imposed" on the system [or that] any given constellation of ... exogenous variables will inevitably lead the system, possibly through a succession of steps, to a state of rest characterized by unchanging prices and production patterns over time ... Once ... we allow for increasing returns, the forces making for continuous changes are *endogenous*—"they are engendered from within the economic system."

Romer (1986, 1987) has recently formalized this argument as it applies to long term growth, showing that growth can become endogenous in the presence of increasing returns. But the idea is of as much (if not more[3]) relevance to the economy's *short-run* dynamics.

The distinction I am drawing between the short- and long-run components of aggregate dynamics is the usual distinction between what can be roughly—and not necessarily in a stochastic sense—referred to as the "stationary" and "nonstationary" components. From a theoretical point of view, models restricted to explain the short- or the long-run component must have reasonable differentiating features that account for stationarity or nonstationarity. Naturally, the sources of increasing returns I rely on to explain short-run fluctuations are different from the ones that have been advanced to explain secular growth. Romer has pinpointed spillovers of knowledge (1986) and increased input specialization (1987) as sources of growth. In this essay I fix the level of knowledge and the degree of specialization and rely on the more standard internal economies of scale or on external economies due to complementarities in firms' output (e.g. Diamond's (1982) "thick market" economies) as causes of cyclical fluctuations. But, as I show, the crucial aggregate consequence that results from any of these short- or long-run sources of increasing returns is the same: the return on investment will typically be an increasing function of the aggregate capital stock—a condition which I call GIRI (Global Increasing Returns on

[3]Although increasing returns give rise to endogenous long-run growth, they are not a necessary condition. If no factors are fixed in the long run, endogenous growth can arise under constant returns (Uzawa (1965)).

Investment). As in Romer's model, this will typically *destabilize* the economy's steady state.

But the fact that increasing returns destabilize the steady state even in a model with short-run sources of scale economies raises the question whether dynamics will not be explosive in this kind of model as well, in which case the idea of a "stationary" component of dynamics is an empty concept and there would be no grounds on which to distinguish short from long-run dynamics. The fact is that differences in the sources of returns to scale do not in themselves allow us to differentiate between short and long-run dynamics. An additional differentiating feature is needed that puts short-run constraints on the economy's growth rate and thus makes for "stationary" endogenous fluctuations *despite* the steady state's instability. In actual economies—with their multiple sectors, their heterogeneous agents, their monetary institutions, etc.—these constraints can arise from various, perhaps quite complex mechanisms. But in this essay I restrict myself to one simple constraining factor which I take as a proxy for possibly more complex mechanisms, namely the presence of resources (infrastructure, developed land, etc.) that are fixed in the short-run. The fixity of these resources result in "congestion" externalities which, as I show, can give rise to external *expansion diseconomies* that put limits on the economy's growth and make for "stationary" dynamics around the economy's unstable steady state (see Howitt and McAfee (1988a)).

I formulate my argument in the general context of dynamic economies in which a "representative agent" maximizes a concave intertemporal "value" function of some state variable ("capital") and of changes therein ("investment"). Special cases of this class of models are basic neoclassical models of savings and of investment: the decentralized version of Ramsey's (1928) model of savings, and the general equilibrium version of the adjustment-cost model of investment first developed by Eisner and Strotz (1963). I show that the crucial determinant of dynamics is whether the return on agents' investment is decreasing or increasing with the aggregate capital stock. In the former case, which arises under the neoclassical assumption of constant returns, the economy has a *saddle path stable* steady state to which it converges. In the latter case (i.e., if GIRI holds), which typically arises under increasing returns, the economy's steady state is *unstable*. Moreover, if external "expansion diseconomies" are introduced to make this a model of short-run fluctuations, I show that the economy is likely to converge to a *stable limit cycle* around this unstable steady state. Fluctuations will persist even in the absence of exogenous shocks.

Besides giving rise to endogenous fluctuations, increasing returns may also result in a *multiplicity of equilibrium paths* that promises further insights into the nature of business cycles. This multiplicity is of two kinds: (1) If GIRI holds only locally and there are regions of increasing and decreasing returns, the economy may have many steady states, which are alternately stable and unstable. But because of "coordination failure," it may get caught in fluctuations around an unstable one. (2) The equilibrium paths that converge to a stable limit cycle are necessarily indeterminate, and can therefore give rise to an *additional* source instability due to "sunspots" (as in Azariadis (1981)). Under increasing returns, exogeneous "animal spirits" can have an important exacerbating effect on economic fluctuations.

Because of this and because nothing precludes more standard external shocks in the above class of models, it would be a mistake to think that endogenous theories of business cycles based on increasing returns necessarily imply "sinusoidal" fluctuations of output around trend. All the reasons for added complexity and for irregularity—from structural change within the system to exogenous disturbances from without—remain valid. But endogeneous theories do suggest that beyond the effects of these disturbances, there may be a more fundamental source of instability in industrial economies that is rooted in the behavior of optimizing agents and in the type of technology in which they operate.

The essay is divided into two parts. (1) The first analyzes two short-run sources of increasing returns: (1.1) external economies due to "thicker" markets, and (1.2) standard internal economies of scale. In both cases I derive conditions under which global increasing returns on investment (GIRI) will hold. (2) The second part makes the argument that business cycle dynamics can become endogenous under increasing returns. I start by discussing two examples of models in which this is true: (2.1) the decentralized Ramsey model of savings, and (2.2) the general equilibrium adjustment-cost model of investment. In both examples I assume that either of the two sources of increasing returns discussed in part 1 is present and gives rise to GIRI. I show that GIRI destabilizes the economy's steady state, and that in the presence of "expansion diseconomies" the economy is likely to converge to a stable limit cycle around it. (2.3) I then extract the essential features of these examples by generalizing the results to a broad class of models. The last section (2.4) discusses the multiplicity of equilibria that can arise in this general class of models, and the resulting role that "animal spirits" and coordination failures may play in the cycle. (3) Finally, I conclude by summarizing the main assumptions and results in the essay. For expository purposes, many results are not proven in the main text. This task is left to the appendix where a general model that encompasses most of the ideas in the essay is treated formally.

1. GLOBAL INCREASING RETURNS ON INVESTMENT

In this part I look at two possible sources of increasing returns and their aggregate implications on the return on investment. Because my focus is on short-run fluctuations, I do not rely on the same sources of increasing returns as are taken to explain long-run growth. Such factors as the level of "knowledge" or input specialization are assumed fixed for my purposes. I describe two ways in which increasing returns can arise which firms can take advantage of in the short-run—namely, external thick-market economies and pure internal economies of scale—and I analyze their aggregate implications on the marginal profitability of capital in a general equilibrium setting.

More formally, I look at the way instantaneous equilibrium in the goods market is determined by the existing capital stock under these assumptions. The implications that will be useful for the study of dynamics in part 2 are summarized in the properties of the representative firm's maximum profit function $\pi(k, K)$, where k and K are the individual and the aggregate capital stocks. Since I am interested in short-run dynamics, I assume for simplicity that the number of firms/industries is fixed and normalize it to one so that, in symmetric equilibrium $K = k$. The main result is that under increasing returns the marginal profitability of capital $\pi_1(k, K)$ will typically rise if all firms expand their capital stock simultaneously. In other words the function

$$\mu(K) \equiv \pi_1(K, K) \tag{1.0.1}$$

is increasing in K ($\mu' > 0$). This property—which I call GIRI (Global Increasing Returns on Investment)—turns out to be the main cause of instability in the class of models I describe in part 2 and is in sharp contrast with the neoclassical case with constant returns where the return on investment falls with the aggregate capital stock ($\mu' < 0$).

1.1. *External Economies*

Although the two above-mentioned sources of increasing returns could arise simultaneously, I analyze them separately. I first consider the possibility of external economies of scale due complementarities in firms' output, and focus primarily on economies due to thicker markets of the sort described by Diamond (1982). In general the situation can be described in terms of the existence of transaction costs which depend negatively on the amount of activity in the market (see Howitt (1985)). The issues can be most simply discussed in the context of a perfectly competitive economy in which all firms produce a quantity y of a single good using a *given* amount of capital k with the concave production function

$$y = F(k), \quad F' > 0, F'' < 0$$

(allowing for an additional "labor" input would not change the results). Each firm incurs in addition the transaction cost $\tau(Y)$ per unit of output sold, which depends negatively on aggregate output Y ($\tau' < 0$). Under symmetry, if we normalize the number of firms to one we get:

$$K = k \quad \text{and} \quad Y = y . \tag{1.1.1}$$

Thus the representative firm's profit function is given by:

$$\pi(k, K) = (1 - \tau(F(K)))F(k).$$

Under what conditions will GIRI arise in this model? If the "real" value of capital does not depreciate over time, the marginal profitability of investment defined in (1.0.1) can be thought of as the rental rate at which the firm's given capital stock *would* have been chosen optimally if firms were allowed to do so. Assuming firms take aggregate output as given, the FOC of this hypothetical maximization problem is[4]:

$$\mu = (1 - \tau(F))F'.$$

Taking (1.1.1) into account, this gives us an expression for μ':

$$\mu' = -\tau' F' + (1-\tau)F''. \tag{1.1.2}$$

The first term is positive by assumption on τ', but the second is negative if any firm is to be in business at all ($\tau < 1$). Thus μ' will be positive and the return on investment will be increasing with the aggregate capital stock as long as the positive effect of the transaction externality is not overcome by a sharply decreasing marginal productivity of capital. This will hold in particular in the limit case of a linear production function ($F'' = 0$).

Generally speaking, GIRI holds under external economies as long as they are not overcome by internal diseconomies, which makes the external and internal sources of scale economies complementary.

1.2. *Internal Economies*

I now turn to the case of internal economies, and to the question whether increasing returns on investment can arise globally (GIRI) if scale economies are only present at the firm level. The presence of increasing returns at the firm level complicates the analysis because it forces us into the domain of general equilibrium analysis with imperfect

---

[4]The second order condition is satisfied because F is concave.

competition and multiple factors of production[5], but the insights make the exercise worthwhile. I analyze the question in the context of the monopolistically competitive economy described by Kyiotaki (1985) and derive conditions under which returns on investment are globally increasing. A more formal treatment of these issues is left to the appendix.

The Setup. Consider an economy with identical firms which can use their given capital endowment k and hire some labor l to produce a quantity y of a differentiated good with the increasing returns production function

$$y = F(k, l), \quad F_k, F_l, F_{kl} > 0, \quad F_{kk}, F_{ll} < 0.$$

Increasing returns to scale roughly means that marginal productivities of capital and labor are greater than their "average" productivities, and therefore that they must be rising with scale "on the average." In this section I strengthen this property by assuming that the two factors' marginal productivities are rising with scale *everywhere* in the region of increasing returns. (This is true in the case of a Cobb-Douglas technology, e.g.). I do this because relying only on the "average" property would yield GIRI only "on the average," and would make the analysis of dynamics in part 2 indeterminate. But there is at least some justification to this assumption if we consider that F is the economy's "representative" production function, and must therefore have "representative" (or "average") properties.

Each firm is a monopolist in the market for its own differentiated product. Under symmetry, firms have the same level of capital and their products enter the representative consumer's utility function in a symmetric way. This implies that they will behave and, in particular, will price their products in an identical manner, so that the aggregate price level in terms of any "representative" product will be unity—which provides us with a handy definition of "real" prices[6].

Each monopolist maximizes its real profits taking the aggregate price level and the demand function for its product as given. It is well known that under these circumstances a monopolist will restrain output enough so as to set the marginal productivity of labor at a markup over the real wage w:

$$pF_l(k, l) = (1+\beta)w, \qquad (1.2.1)$$

[5]Unless we are willing to make the uncomfortable assumption that the marginal productivity of capital is increasing.

[6]This is only valid if the differentiated goods in question are the only goods that enter the utility function. See the appendix for more details.

where p is the real price the firm sets for its product. The markup factor $1+\beta = 1/(1-\eta^{-1})$ is greater the smaller is the (absolute value) elasticity of demand for its product η. From here on I assume a locally constant elasticity $\eta > 1$ (and markup $\beta > 0$)[7].

Similarly, *if* it were allowed to choose its capital endowment k, the monopolist would set the marginal productivity of capital at the same markup over the the rental rate μ on capital:

$$pF_k(k, l) = (1+\beta)\mu. \tag{1.2.2}$$

In fact I assume the firm's capital stock is given, but equation (1.2.2) is still useful because it gives us an expression for the "shadow" rental rate on capital μ at which the firm's capital stock would have been optimal. This is none other than the marginal profitability of capital defined in (1.0.1) and in which we are interested.

The question about GIRI is whether a symmetric increase in the aggregate level of capital increases its marginal profitability μ. If all firms expand their capital stock by the same amount they will remain symmetric and charge the same real price of unity ($p = 1$). Thus the FOCs[8] (1.2.1)-(1.2.2) become:

$$F_l(K, L) = (1+\beta)w; \tag{1.2.3}$$

$$F_k(K, L) = (1+\beta)\mu, \tag{1.2.4}$$

where k and l have been replaced by their aggregates K and L after the usual normalization of the number of firms to one. This being a system of two equations in four unknowns (K, L, w, and μ), I must add one more equation before I turn to the dependance of μ on K. Naturally, the additional relationship is the positive wage/employment relationship

$$w = w(L)$$

usually associated with "labor supply," although I will not limit myself to this interpretation. Given this relationship $d\mu/dK$ is obtained by totally differentiating (1.2.3)-(1.2.4):

$$\frac{d\mu}{dK} = (1+\beta)^{-1}\left(F_{kk} + \frac{F_{kl}^2}{(1+\beta)w'-F_{ll}}\right). \tag{1.2.5}$$

I show that $d\mu/dK > 0$ and returns on investment are increasing in the aggregate if w' is small enough, i.e., if the aggregate "labor supply" curve "looks" relatively flat *in the short-run* (since I am interested in short-run fluctuations), which seems to be empirically valid.

---

[7]This will be true with an underlying CES utility function behind the demand for differentiated products.

[8]The SOC to profit maximization (with respect to both k and l) is also assumed to hold. It is $\xi_k + \xi_l < 1+\beta$ in the CES case.

The Labor Market. To build some intuition for this condition, consider its polar opposite, i.e., a vertical labor supply curve ($w' = \infty$). Suppose workers supply all their labor endowment inelastically and wages clear the labor market, so that the economy is always at full employment. A symmetric increase in k raises F_l in (1.2.3), but since aggregate employment cannot increase, it is the real wage that will adjust to match the higher marginal productivity of labor. Since employment remains unchanged, the rise in k depresses F_k (unless $F_{kk} > 0$, which would be a rather controversial way of generating GIRI) and the marginal profitability of capital μ in (1.2.4) *falls* due to its decreasing marginal productivity. What is happening is that, because of the inelasticity of labor supply, firms will not be able to increase employment simultaneously in response to the rise in their capital stock so as to take *effective* advantage of increasing returns. The aggregate constraint on total employment prevents internal economies of scale from giving rise to aggregate economies.

But this situation is far from realistic. Aggregate employment does in fact vary much over the cycle, and real wages do not.[9] Based on this fact, in the following discussion I fix the real wage equal to a constant w^* for simplicity ($w' = 0$) and assume that in fact it will not change so drastically as to reverse the conclusions. Of course, this is consistent with a wide range of interpretations that have been given to the empirical short-run wage/employment relationship. It would arise, for example, under (1) the standard Keynesian short-run flat labor supply curve; or (2) an efficiency wage story, where the wage rate has a positive effect on workers' morale and productivity (or, more generally, plays another role in addition to its role as compensation for labor services) and is set optimally at the "efficiency" level w^* that maximizes worker effort per dollar paid in wages;[10] or (3) a neoclassical account, where there exists a "home" production technology with which a unit of labor produces w^* units of "leisure" that provide the same utility as consuming differentiated products.[11] (In the appendix I treat the "efficiency wages" case, but the argument does not depend on this specific interpretation.)

Global Increasing Returns. Once we allow firms to take advantage of scale economies by increasing employment in response to a rise in their capital stock, it becomes clear why

---

[9] See, e.g., Geary and Kennan (1982) and Bils (1985).
[10] This argument is due to Solow (1979) and is developed in more detail in the appendix. For a more general survey of efficiency wage arguments see Katz (1987).
[11] The resulting instantaneous utility function is $U(C\text{-}w^*L)$, where L designates the employment level. The implied strong degree of substitutability between consumption and leisure may be unappealing. But, given the small cyclical variability of real wages, this substitutability is necessary for a procyclical consumption pattern if preferences are time separable (see Barro and King (1984)).

economies of scale at the individual level can be transmitted to the aggregate economy. Recall that I assume marginal productivities to be increasing with scale. Now suppose K is symmetrically increased. Labor's marginal productivity F_l in (1.2.3) will rise above the "rigid" wage w^*, and firms will find it profitable to increase employment up to the point where F_l again equals w^*. In fact firms will increase employment *more than proportionally* in response to the rise in K, because if they were to increase it just proportionally they would be precisely increasing the scale of their operations and, by what was assumed about technology, they would end up with a marginal productivity of labor that is higher than it was initially, i.e., that is higher than w^*. This implies that F_k will rise, since, again by my technological assumption, it would have risen with the scale of operations even if firms had increased employment only proportionally. Thus, by (1.2.4), the marginal profitability of capital μ, which is proportional to F_k , will also rise, and GIRI holds. In sum, firms will increase employment more than proportionally in response to an increase in K, thus taking more than full advantage of increasing returns and increasing the marginal profitability of their capital.

It is important to emphasize the *pecuniary externality* underlying this mechanism, which is buried under the fixing of firms' real price p in (1.2.1)-(1.2.2) equal to one along the expansion process. To see what would have happened in the absence of this externality, consider the effect on marginal profitability μ of a change in a firm's *own* capital stock k, aggregates unchanged. The result (which I denote by $d\mu/dk$ with a slight misuse of notation) can be inferred from the individual FOCs (1.2.1)-(1.2.2) and is given by:

$$\frac{d\mu}{dk} = - (1+\beta)^{-1} \frac{pF_k}{k}\left(\frac{(1+\beta) - (\xi_k+\xi_l)}{(1+\beta) - \xi_l}\right) \tag{1.2.6}$$

under constant elasticities. Now, the only reasonable sign for $d\mu/dk$ is negative, because otherwise the firm would choose to invest infinitely whenever it has a chance. This is in fact guaranteed by the SOC to profit maximization (with respect to both k and l):

$$\xi_k + \xi_l < 1+\beta.$$

The difference between this and the aggregate experiment is that here output expansion has a negative effect on prices, which *must*, by the SOC, overcome the positive effect of scale economies on μ.

If firms expand their capital stock simultaneously this negative effect disappears because every firm's demand curve shifts up as it grows by the exact amount that is required to prevent its price from falling (this is what allowed me to set $p = 1$ in (1.2.3)-(1.2.4)). The rise in demand is in turn due to the "Say's law" case of supply creating its own demand through the growth in the income of consumer-owners. (The fact that these effects were only implicit in my discussion is the price I had to pay for simplicity of exposition, but a more explicit treatment can be found in the appendix). Under constant elasticities, the effect of the aggregate experiment on μ is given by:

$$\frac{d\mu}{dK} = - (1+\beta)^{-1} \frac{pF_k}{k} \left(\frac{1 - (\xi_k + \xi_l)}{1 - \xi_l} \right)$$

which differs from (1.2.6) only insofar as β is absent from the expression in parentheses. In this sense, firms react to a symmetric rise in their capital stock *as if* they were price takers ($\beta = 0$) reacting to a rise in their own capital stock, so that no negative effect on prices offsets the positive effect of increasing returns ($\xi_k+\xi_l > 1$) and μ rises.

I sum, the above exercise shows that if we assume (1) that the marginal productivities of capital and labor are increasing with scale, and (2) that there is some degree of short-run "real wage rigidity," then internal increasing returns at the firm level give rise to global increasing returns on investment (GIRI).

2. ENDOGENOUS BUSINESS CYCLES

Having found plausible conditions under which scale economies—either internal or external—can give rise to GIRI, I turn to their implications on business cycle dynamics. I consider a general class of dynamic economies and show that GIRI will typically destabilize their steady state. In the presence of external "diseconomies of expansion" (as in Howitt and McAfee (1988a)) in the short-run, I show that equilibrium paths are likely to converge to a stable limit cycle. Fluctuations will persist endogenously, even in the absence of external shocks. I first discuss in detail two examples of special cases of the above class that are particularly popular with exogenous business cycle theorists: the decentralized Ramsey model of savings, and the adjustment cost model of investment. Both are treated more rigorously in the appendix.

2.1. *A Savings Cycle*

In this and the next example the economy has two kinds of agents—households and firms. Throughout I maintain the symmetry assumption and normalize the number of both consumers and firms to one, so that aggregate output, consumption, investment, capital, employment, and dividends are equal to their "representative agent" values:

$$Y = y;\quad C = c;\quad I = i;\quad K = k;\quad L = l;\quad D = d. \tag{2.1.1}$$

In both examples the productive sector can be indifferently modelled after either of the two models of increasing returns in part 1 (in the appendix I treat the internal economies case). I use three main results that can be derived under either model: (1) Firms' profit function $\pi(k, K)$ satisfies GIRI:

$$\mu(K) \equiv \pi_1(K, K),\ \mu' > 0; \tag{2.1.2}$$

This is to be contrasted with the constant returns case where $\mu' < 0$. (2) Aggregate output is determined by the aggregate capital stock:

$$Y = Y(K),\ Y'>0; \tag{2.1.3}$$

(3) The industry can be treated *as if* it were producing a single "representative" good that can be used for either consumption or capital accumulation. This is true by assumption for the first model, and because of symmetry in the second, if we assume that the differentiated products enter both the utility and production function in the same way (see appendix). In particular we can write the goods market equilibrium equation:

$$Y = C + I. \tag{2.1.4}$$

Firm and Household Behavior. This first example is that of the decentralized version of the familiar Ramsey model of savings. Firms invest at the net rate

$$i = dk/dt. \tag{2.1.5}$$

without any capital adjustment costs[12]. They maximize the present discounted value of profits at the market interest rate r, taking aggregates as given. It follows (by the FOC) that r must equal the marginal profitability of capital:

$$r = \mu(K). \tag{2.1.6}$$

Households supply their labor endowment inelastically[13] and have subjective discount rate $\rho > 0$. They maximize

$$\int_t^{\infty} U(c(s))e^{-\rho(s-t)}ds \tag{2.1.7}$$

[12]In the class of models I am considering, no nonnegativity constraint is imposed on investment, but there are two reasons why this is not very restrictive: (1) Allowing for negative net investment does not at all require investment decisions to be reversible. Net disinvestment is perfectly compatible with positive gross investment; (2) Even if depreciation is weak, the nonegativity constraint can be approximated by the presence of capital adjustment costs, which are incorporated in the next example.

[13]This is compatible with two of the labor market explanations in section 1.2 for "real wage rigidity," but not with the "neoclassical" account. However the model could easily be modified without major changes in the results to incorporate this third explanation by changing the instantaneous utility function to $U(c - w^*l)$.

subject to their wealth constraint, where U is a concave increasing function ($U' > 0$, $U'' < 0$). As is well known, under perfect foresight the optimal consumption path will satisfy the FOC:

$$\frac{\dot{c}}{c} = \sigma(c)(r - \rho), \quad \sigma(c) = -\frac{U'(c)}{U''(c)}\frac{1}{c} > 0, \tag{2.1.8}$$

where $1/\sigma$ is the rate of intertemporal substitution. Consumption is increasing over time if $r > \mu$, and decreasing otherwise.

Dynamics. Equilibrium dynamics are two dimensional and can be studied in (K, C)-space. Taking the normalization (2.1.1) into account, capital dynamics are given by investment in the aggregate counterpart to (2.1.5), which by goods market equilibrium (2.1.4) is equal to savings $Y(K) - C$. Consumption dynamics are driven by the interest rate in aggregate counterpart to (2.1.8), and the latter is a function of K by (2.1.6). Thus, dynamics are governed by the system:

$$\begin{aligned} \dot{K} &= Y(K) - C \\ \dot{C} &= \sigma(C)\, C\, (\, \mu(K) - \rho\,). \end{aligned} \tag{2.1.9}$$

The *phase portrait* is given in Figure 1. The capital stock remains unchanged ($dK/dt = 0$) whenever savings are zero, i.e., $C = Y(K)$. Consumption remains unchanged ($dC/dt = 0$) if the rate of interest equals the subjective discount rate[14], i.e., when the capital stock is at the "modified golden rule" level K^* defined by:

$$\mu(K^*) = \rho.$$

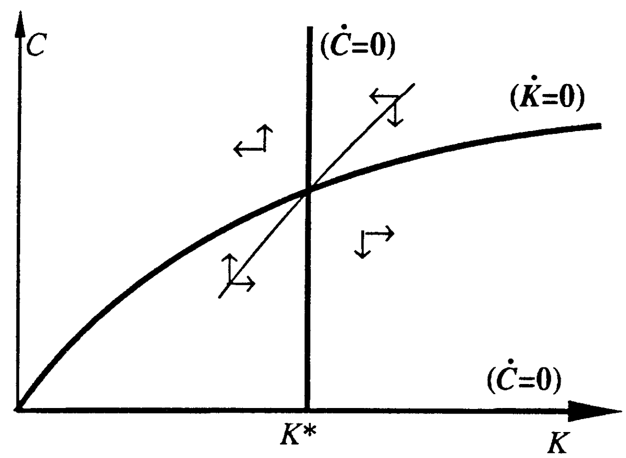
Thus, besides the degenerate case when $K = C = 0$, the economy's *steady state* is given by the "modified golden rule" level of capital K^* (which is assumed to exist) and a corresponding level of consumption that leaves nothing to savings: $C^* = Y(K^*)$.

The main determinant of dynamics is whether μ is increasing or decreasing with K, i.e., whether GIRI holds or not. As illustrated in figure 1, the dynamics around steady state are *saddle path stable* in the neoclassical case where $\mu' < 0$, but they are *spiralling* if $\mu' > 0$. To see why this is so note that if $\mu' < 0$, r is higher than ρ when $K < K^*$, and lower afterwards; but if $\mu' > 0$, the reverse is true. Thus by the FOC (2.1.8), under decreasing returns the consumption path will be rising if K is below its steady state value, while under increasing returns it will be falling. This implies that if $\mu' < 0$ there must exist a saddle path where households start with positive savings when $K < K^*$, increase consumption as capital accumulates, and reduce savings down to zero as the steady state K^* is approached.

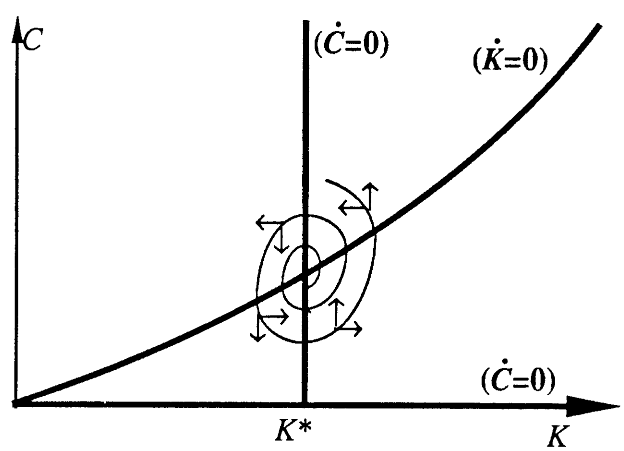
---

[14] As long as $\sigma(0)$ is finite, consumption will also remain unchanged when it is nil, so that no path that starts with positive consumption will lead to a negative level of consumption.

But if $\mu' > 0$, household behavior becomes in a sense perverse because they will be following a path of *falling* consumption and rising savings when $K < K^*$, so that K^* will necessarily be reached with positive savings and the economy will overshoot its steady state. Beyond K^*, the consumption path starts rising and tends to reverse the growth, but by a similar argument the economy will also undershoot K^* in the following contraction, and will be involved in repeated cycles thereafter.



Dynamics with $\mu' < 0$



Dynamics with $\mu' > 0$

FIGURE 1

The question arises whether these cycles will converge to steady state or not—in other words whether the steady state is stable or not under GIRI. The system's eigenvalues at (K^*, C^*) are given by:

$$\lambda = 1/2(Y' \pm \sqrt{Y'^2 - 4\sigma C\mu'})|_{(K^*,C^*)} \qquad (2.1.10)$$

It can be seen that their real parts are positive if $\mu'(K^*) > 0$[15], and therefore that the steady state is *unstable* under GIRI. Mathematically speaking, if $\mu' > 0$ then $Re(\lambda)$ has the sign of $Y'(K^*)$, and in this sense it is the fact that $Y' > 0$ that is responsible for instability under GIRI. This corresponds to the intuition that income, being an increasing function of the capital stock, tends to destabilize the equilibrium by reinforcing the economy's momentum in the direction in which it is geared. For if savings are positive and output is expanding, income will increase and contribute to even higher savings. If savings are negative and output contracting, income will fall and push savings further down. If, on the contrary, income were a *decreasing* function of the capital stock, it would always tend to decrease the economy's momentum and drive it in the direction opposite to the one in which it is geared, thus stabilizing the steady state.

A Business Cycle. Under GIRI the economy will be fluctuating around an unstable steady state. The next question that arises is whether these fluctuations will be explosive, or whether they will converge to some *stable limit cycle*[16] where the economy undergoes periodic cyclical fluctuations around steady state, as illustrated in figure 2. Although the rigorous mathematical treatment of this question must be left to the appendix, the results are summarized here. In order to find conditions under which dynamics are not explosive, I need to explore the possibility of there being factors that put constraints on the economy's growth rate. The model could be complicated in many directions to achieve this effect, but I only explore here the simplest way in which such constraints can be introduced, taking them as proxies for more complex mechanisms in actual economies.

I take advantage of the insight that it is the fact that Y rises with K that destabilizes the steady state, and I look for factors—which I identify with short-run "congestion" externalities in what follows[17]— that constrain $Y'(K)$ without affecting the model's other features. For now suppose that we have identified a parameter ϕ that affects Y' in such a way that $Y'(K^*)$ varies from negative to positive as ϕ is increased (of course, the only reasonable values of ϕ are the ones corresponding to $Y' > 0$). By my discussion of

---

[15]The eigenvalues' real parts are of opposite signs if $\mu'(K^*) < 0$, which confirms saddle path stability.
[16]For a rigorous definition of a limit cycle see Hirsch and Smale (1974), p. 250.
[17]This interpretation was pointed out to me by Peter Howitt.

stability, this implies that the eigenvalues' real parts in (2.1.10) will also vary from negative to positive as ϕ increases, and the equilibrium will change from stable to unstable. It is said that a *Hopf bifurcation* occurs at the *critical value* ϕ * of the parameter at which both eigenvalues cross the imaginary axis. Hopf's theorem (see appendix) states that a limit cycle will appear in a neighborhood of ϕ on one side of the critical value ϕ *.

The side on which the limit cycle appears—i.e., whether it appears when Y' is positive or negative—depends on a certain stability condition which is calculated in the appendix. If returns on investment are increasing at a decreasing rate ($\mu'' < 0$), the condition requires

$$\sigma < \frac{1}{C^*\mu'(K^*)}$$

in order for the limit cycle to appear on the "reasonable" side of the critical value where $Y' > 0$. The rate of intertemporal substitution $1/\sigma$ should not be too small, and the requirement is weaker the weaker is the degree of GIRI (as measured by $\mu'(K^*)$). But in the case where GIRI is extreme and $\mu'' > 0$, the requirement is:

$$\frac{1}{C^*\mu'(K^*)} < \sigma.$$

$1/\sigma$ should not be too large, and the requirement is stronger the smaller is $\mu'(K^*)$.

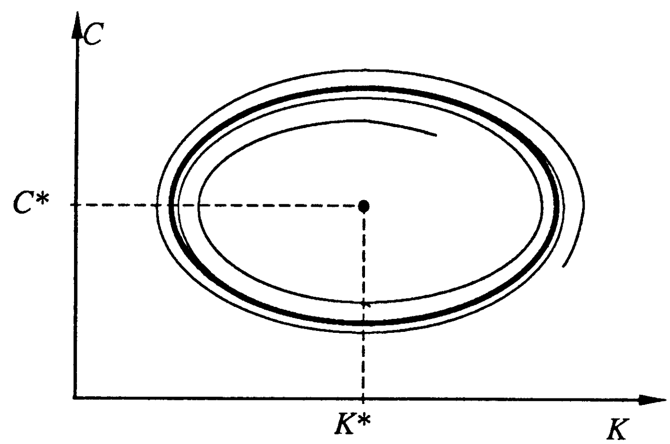


FIGURE 2

Congestion. I must now come to the promised "congestion" interpretation of the parameter ϕ. Congestion is introduced into the model by assuming that there are some resources used by all firms—such as infrastructure, land, etc.—that are fixed in the short-run (since I am interested in short-run fluctuations). As a result an increase in aggregate activity has a negative effect on individual firms' production function. To capture this idea I make the production function F depend negatively on aggregate output:

$$F(k, l; Y) = f(k, l) - vY, \quad v > 0,$$

where v is a measure of the degree of congestion. Solving under symmetry for Y as a function of K (after taking the normalization (2.1.6) into account and the fact that aggregate employment L will in general be determined by K[18]) we find that

$$Y(K; \phi) = \phi f(K, L(K)), \quad \phi \equiv (1+v)^{-1}. \tag{2.1.11}$$

This makes Y' depend on ϕ in the required manner.

But now I must also make sure that, as assumed above, the model's other important features are not affected by ϕ, and in particular that the function $\mu(K)$ is not. This is true in the internal economies model of section 1.2 by virtue of the fact that Y is assumed to enter additively in $F(k, l; Y)$[19]. It is not in the external economies model of section 1.1 because the transaction cost τ which determines μ is affected by Y and therefore by ϕ. But if, in the *thought experiment* of varying ϕ, we allow the transaction externality to increase with congestion by replacing the transaction cost $\tau(Y)$ by $T(Y; \phi) = \tau(Y/\phi)$, μ would no longer depend on ϕ.

In this example I have examined the effect of increasing returns on the basic neoclassical model of savings. I have shown that, contrary to the neoclassical presumption of saddle path stability, GIRI destabilizes the economy's steady state. Moreover, if (1) there is some degree of "congestion" in the economy, and if (2) the rate of intertemporal substitution is in a certain range, then the economy will converge to a stable limit cycle. Fluctuations will persist endogenously, even in the absence of external shocks.

2.2. *An Investment Cycle*

In the previous example, capital dynamics originated primarily in households' *savings* behavior while firms' *investment* behavior remained passive. This is because households were maximizing a *concave* utility function that gave rise to a problem of intertemporal substitution, while firms maximized a *linear* function of investment that made investment perfectly substitutable over time. But firm investment in itself can also give rise to interesting dynamics once we relax this linearity assumption. In this example I allow for

---

[18]This is trivially true in the model in section 1.1, and follows from (1.2.3) in the second model.

[19]The polar opposite where Y enters multiplicatively would not work. In general v must measure the sensitivity of F to the additive component of the effect of Y.

convex capital adjustment costs due to the fact that it is difficult for firms to change their capital stock instantly to any desired level (see Eisner and Strotz (1963), Lucas (1967a, b), Gould (1968), Treadway (1969)). In order to focus on intertemporal substitution due to adjustment costs, I "neutralize" households' intertemporal activity by assuming their utility function to be linear. However, the results do not depend on this simplifying assumption: In the appendix I treat the case where both households and firms maximize concave functions, and the results turn out to be "weighted averages" of the results in this and the previous section. The "weights" in question depend on the curvatures of the functions maximized by households and firms.

The Economy. Households in this economy face the same utility maximization problem (2.1.7) as before, but now I assume for simplicity that U is linear ($U'' = 0$). Since this means that consumption is perfectly substitutable over time, the FOC implies that the market interest rate must always equal their subjective discount rate:

$$r = \rho. \tag{2.2.1}$$

Firms face capital adjustment costs. In order to increase their capital stock at the *net* rate

$$i = dk/dt. \tag{2.2.2}$$

they must expend an amount

$$g = g(i; \xi K), \quad g_i > 0, g_{ii} > 0, g_{iK} > 0, \xi > 0. \tag{2.2.3}$$

The convexity of the adjustment cost g in i guarantees the concavity and regularity (in terms of SOCs) of the profit maximization problem. Note that g is also assumed to depend on the aggregate capital stock K with a "sensitivity" parameter ξ. This is done in anticipation of the fact that the presence of external "expansion diseconomies" will be needed, as in the previous example, to prove the existence of a limit cycle. This externality can again be interpreted as arising from the presence of fixed resources in the short-run, which are used up in the capital accumulation process. As the aggregate capital stock increases, these resources become increasingly scarce and there results an increase in the marginal cost of investment ($g_{iK} > 0$). Under this interpretation, ξ is a measure of congestion.

Firms maximize the present discounted value of profits

$$\int_t^\infty \left[\pi(k, K) - g(i, \xi K)\right] e^{-\int_t^s r(u)du} ds \tag{2.2.4}$$

at the market interest rate r, taking aggregates as given. Under perfect foresight, the FOC for this problem is:

$$\frac{di}{dt} = \frac{g_i}{g_{ii}}\left[r - \left(\frac{\mu}{g_i} + \frac{\xi g_{iK}}{g_i} I \right) \right]. \tag{2.2.5}$$

If we define the return on investment adjusted for present installation costs as:

$$\overset{a}{\mu}(K, I; \xi) \equiv \frac{\mu(K)}{g_i(I, \xi K)}, \tag{2.2.6}$$

it can be seen that investment is increasing over time when the interest rate is greater than the "adjusted" return on investment plus a term accounting for expected changes in installation costs, and decreasing otherwise. Although the presence of congestion introduces a negative effect of K on μ^a, I will assume whenever needed that this does not overcome the positive effect due to GIRI so that μ^a is increasing in K (I will come back to this assumption).

The economy's equilibrium dynamics are given by the aggregate counterparts to (2.2.2) and (2.2.5), taking into account the normalization (2.1.1) and the determination of market interest rates in (2.2.1):

$$\dot{K} = I$$
$$\dot{I} = \frac{g_i(I, \xi K)}{g_{ii}(I, \xi K)}\left[\rho - \left(\overset{a}{\mu}(K, I; \xi) + \frac{\xi g_{iK}(I, \xi K)}{g_i(I, \xi K)} I \right) \right]. \tag{2.2.7}$$

Business Cycle Dynamics. The dynamics can be studied in (K, I)-space. A *steady state* (K^*, I^*) for the economy must have zero investment and the adjusted return on investment must equal the interest rate:

$$I^* = 0 \quad \text{and} \quad \mu^a(K^*, 0; \xi) = \rho.$$

Assuming such a K^* exists[20], the eigenvalues at (K^*, I^*) are given by:

$$\lambda = \frac{1}{2}\left[\left(\rho - \frac{\xi g_{iK}}{g_{ii}} \right) \pm \sqrt{ \left(\rho - \frac{\xi g_{iK}}{g_{ii}} \right)^2 - 4 \frac{g_i}{g_{ii}} \mu_K^a } \right]_{(K^*, I^*)}. \tag{2.2.8}$$

First consider the case when there is no congestion ($\xi = 0$). As in the previous example, under the neoclassical presumption that $\mu' < 0$, the equilibrium is saddle path stable ($\lambda_1 \lambda_2 < 0$), but under GIRI ($\mu' > 0$) it is totally unstable ($Re(\lambda_1), Re(\lambda_2) > 0$) and dynamic paths are spiralling outwards from (K^*, I^*).

[20] If K^* exists, then it is unique in the range where μ^a is increasing in K.

Now if we introduce congestion under GIRI, equilibrium paths can again be shown to be spiralling into a limit cycle around steady state. Assume for simplicity that g is of the form:

$$g(i, \xi K) = \gamma(i) + (\xi K)i, \quad \gamma, \gamma' > 0, \tag{2.2.9}$$

so that $g_{iK} = 1$. As ξ rises above zero, the term $(\rho - \xi g_{iK}/g_{ii})$ in (2.2.8) falls and, since $\mu^a{}_K$ is assumed positive, the eigenvalues cross the imaginary axis at the critical value:

$$\xi^* = \rho g_{ii}(I^*, K^*). \tag{2.2.10}$$

The condition that μ^a should still be increasing in K when $\xi = \xi^*$ guarantees that the eigenvalues remain imaginary. By (2.2.6), (2.2.9) and (2.2.8) it is equivalent to:

$$\rho^2\gamma''(I^*) < \mu'(K^*). \tag{2.2.11}$$

One way to view this condition is that it requires adjustment costs not be very convex for low levels of net investment (i.e., γ'' must be small at $I^* = 0$)[21].

In this case a Hopf bifurcation occurs at the critical value ξ^*, and a limit cycle will appear in a neighborhood on one side of ξ^*. In order for the limit cycle to appear at the reasonable side $\xi < \xi^*$ where congestion is not excessive, a stability condition must be satisfied. As shown in the appendix, the condition is:

$$\frac{\gamma'''(0)}{\gamma''(0)} \notin [0, 1/2]. \tag{2.2.12}$$

If the marginal cost of adjustment is increasing at a decreasing rate $(\gamma'''(0) < 0)$, the condition is always satisfied. But if $\gamma'''(0) < 0$, it is only satisfied when adjustment costs are not very convex for low investment levels $(\gamma''(0) < 2\gamma'''(0))$, a type of condition already needed in (2.2.11).

This example gives a second popular model—namely the adjustment cost model of investment—where GIRI has an essentially destabilizing effect. Moreover, it was shown that if (1) there are external expansion diseconomies, and (2) adjustment costs are not excessively convex at the origin, the economy will converge to a stable limit cycle around steady state.

---

[21]This is not the same as requiring adjustment costs not to be very convex for zero *gross* investment, which would be problematic in the presence of irreversibility.

2.3. *A Class of Models*

It is possible to generalize the results obtained in the above two examples to a fairly broad class of models that incorporate their essential features, namely (1) agents' maximization of a concave intertemporal "value" function of some state variable ("capital") and changes therein ("investment"); (2) the fact that the "return" on agents' "investment" rises with a simultaneous expansion in the aggregate "capital" stock (GIRI); and (3) the presence of external expansion diseconomies. I present in what follows the simplest general class models that incorporate these features and show how they naturally give rise to endogenous fluctuations. The discussion can be generalized to a broader class with multiple state variables, many kinds of "intertemporally active" representative agents (as in the model treated in the appendix), etc.

Assumptions. At any instant in time, each agent owns an amount k of "capital" which he can increase by "investing" at the *net* rate $i = dk/dt$. He discounts at rate ρ an instantaneous "value function"

$$V = V(R, i, E), \quad V_R > 0, \quad V_i < 0, \quad V_{iE} < 0, \tag{2.3.1}$$

which depends on the revenue R he gets from owning capital, his investment level i, and an aggregate "externality" E. It is important to realize that it is the assumed time separability of the problem that will allow us to draw strong conclusions about dynamics.

V depends positively on revenue and negatively on the investment rate (foregone revenue). In turn, the agent's revenue increases with his capital stock k, and also depends on the aggregate capital stock K:

$$R = R(k, K), \quad R_k > 0. \tag{2.3.2}$$

We can formulate the GIRI hypothesis by normalizing the number of agents to one, so that $K = k$ in symmetric equilibrium, and defining the return on investment as a function of K:

$$\mu(K) \equiv R_k(K, K).$$

A symmetric rise in K increases the return on investment (GIRI holds) if $\mu' > 0$.

Finally, V is affected by an aggregate "externality" E, which is itself a function of the aggregate capital stock. The important characteristic I assume about this external effect is that it is gives rise to what Howitt and McAfee (1988a) call "expansion diseconomies": the marginal "cost" of investment increases with the externality ($V_{iE} < 0$, as indicated in (2.3.1)). A parameter ξ is introduced in order to discuss changes in the strength of the externality, a higher level of ξ indicating a greater marginal contribution of K to the externality ($E_{K\xi} > 0$):

$$E = E(K; \xi), \quad E_{K\xi} > 0. \tag{2.3.3}$$

The meaningfulness of variations in such a parameter is an assumption in itself. In the examples ξ is interpreted as measuring "congestion" due to the presence of resources that are fixed in the short-run.

Each agent chooses a path of investment $\{i(t)\}$ that maximizes the discounted present value of V:

$$\int_t^{\infty} V(R(k(s), K(s)), i(s), E(K(s); \xi))e^{-\rho(s-t)}ds, \tag{2.3.4}$$

subject to $dk/dt = i$, taking the path of the aggregate capital stock $\{K(t)\}$ as given. In order for this maximization problem to be well behaved (in terms of SOCs), the integrand in (2.3.4) is assumed concave in (k, i). In particular:

$$V_R R_{kk} + V_{RR} R_k^2 < 0, \quad \text{and} \quad V_{ii} < 0. \tag{2.3.5}$$

Before I turn to the analysis of dynamics, I must make clear why the two examples given above are special cases of this class of models.

Examples. In order to put the first example's utility maximization problem in the form (2.3.4) I must specify the household's budget constraint and its dependence on aggregates, and then substitute it for "consumption" in the utility function. Without loss of generality[22], assume that households own the labor and the capital stock in the economy and rent them to firms at the wage w and interest rate r. They also own the firms from which they earn dividends d. Their intertemporal budget constraint is therefore:

$$rk + wl + d = c + i.$$

Households take noninterest income $(wl + d)$ as a given that depends on aggregate conditions. Under the normalization (2.1.1), it is equal to the *aggregate* noninterest income in the economy:

$$wl + d = Y - rK.$$

Replacing this term in the budget constraint, and taking into account the dependence of Y and r on K in (2.1.3) and (2.1.6), and of Y on the parameter ϕ in (2.1.11), we get:

$$\mu(K)k + [Y(K; \phi) - \mu(K)K] = c + i.$$

At this stage it is easy to show that this model is a special case of the general class (2.3.4). Just substitute the budget constraint (2.1.11) into the utility function (2.1.7), and the maximization problem becomes:

$$\underset{\{i(s)\}}{Max} \int_t^{\infty} U(\mu(K)k - i + [Y(K; \phi) - \mu(K)K])e^{-\rho(s-t)}ds$$

subject to $i = dk/dt$. It is clearly of the required form, once we write:

[22]The alternative that firms own the capital stock is treated in the appendix and yields identical results.

$$V(R, i, E) = U(R - i - E),$$
$$\text{where} \quad R(k, K) = \mu(K)k, \quad E(K; \xi) = -[Y(K; -\xi) - \mu(K)K], \quad \xi = -\phi.$$

In this example the externality E arises because "noninterest income" depends on aggregate conditions. All the assumptions about V, R, and E can be verified. In particular the problem's concavity comes from the the utility function's concavity, and the GIRI property of R ($\mu' > 0$) comes ultimately from increasing returns.

Putting the second example's profit maximization problem into the required form is more straightforward. Taking the determination of interest rates (2.2.1) and the dependance of adjustment costs on the "congestion" parameter in (2.2.9) into account, the present value of profits (2.2.4) to be maximized becomes:

$$\int_t^{\infty} \left[\pi(k, K) - g(i, \xi K)\right] e^{-\rho(s-t)} ds.$$

This is a special case of (2.3.4) with

$$V(R, i, E) = R - g(i, E),$$
$$\text{where} \quad R(k, K) = \pi(k, K), \quad E(K; \xi) = \xi K.$$

Again all assumptions about V, R, and E can be verified. In particular the problem's concavity comes from the profit function's concavity in k and the adjustment costs' convexity in i. The GIRI property of R ($\mu' > 0$) comes unltimately from increasing returns.

Of course these are not the only economic examples of models of the form (2.3.4), neither should the state variable k necessarily be interpreted as the "capital stock." The model by Howitt and McAfee (1988a) is also a special case with external economies, where *"k"* represents *employment* and congestion arises from the fact that hiring costs rise with aggregate employment (finding new recruits becomes costlier when the pool of unemployed shrinks). Diamond and Fudenberg (1987) also prove the existence of a limit cycle in a formally similar (but not identical) model of search where producers must keep their goods in inventory while searching for a trading partner. The state variable *"k"* is the level of *inventories* and the equivalent of "global increasing returns" arises from the fact that the arrival rate of trading partners—i.e., the "expected return" on inventory investment "μ"—increases with the number of people searching, and therefore with the aggregate level of inventories *"K."* Finally "expansion diseconomies" arise in their model because the rate at which inventories are depleted equals the rate of sales, which naturally rises with the arrival rate $\mu(K)$, and therefore with the aggregate level of inventories K. This is a different mechanism than the congestion externality used in this essay to generate expansion diseconomies, and could arise in the class of models (2.3.5) if we introduce *convex capital depreciation*. The point should again be stressed, that congestion is only taken here as a proxy for the various other possible sources of expansion diseconomies.

Dynamics. The FOC and transversality condition to the value maximization problem (2.3.4) give us the equilibrium conditions once we take the symmetry and normalization assumptions ($K = k, I = i$) into account (the SOC is satisfied because of the "value" function's assumed concavity). As in the previous example, define the "adjusted" (in terms of "value") return on investment as:

$$\mu^{a}(K, I; \xi) \equiv \frac{\mu V_R}{-V_i}. \tag{2.3.6}$$

Then equilibrium dynamics in (K, I)-space are governed by the system:

$$\dot{K} = I$$
$$\dot{I} = \frac{V_i}{V_{ii}}\left\{\rho - \mu^{a} - \frac{1}{V_i}\left[V_{iR}\left(R_k + R_K\right) + V_{iE}E_K\right]I\right\}. \tag{2.3.7}$$

Investment is increasing over time when the interest rate is greater than the "adjusted" return on investment plus a term accounting for expected changes in the "cost" of investment, and decreasing otherwise. The transversality condition is:

$$\lim_{t\to\infty} KV_i e^{-\rho t} = 0. \tag{2.3.8}$$

At a steady state (K^*, I^*) investment must be zero and the adjusted return on investment must equal the discount rate:

$$I^* = 0 \quad \text{and} \quad \mu^a(K^*, 0) = \rho.$$

Assuming a steady state exists[23], the system eigenvalues at (K^*, I^*) are given by:

$$\lambda = \frac{1}{2}\left(B \pm \sqrt{B^2 - 4D}\right),$$

$$\text{where } B = \left(\rho - \frac{V_{iR}}{V_{ii}}R_K\right) - \frac{V_{iE}}{V_{ii}}E_K \quad \text{and} \quad D = \frac{V_i}{V_{ii}}\mu^a_K$$

I assume that at (K^*, I^*) the "adjustment" to the return on investment μ^a does not reverse its dependence on K, i.e., that $\mu^a{}_K > 0$ under GIRI. In this case D has the sign of μ'. This was true in the first example simply because the structure of the model made $\mu^a = \mu$, but in the second example this assumption required that adjustment costs not be too convex at the origin. But, as long as $\mu'(K^*)$ does not go to zero with ρ, it can be shown that a general sufficient condition for the condition to hold is that the discount factor ρ should not be too large.

---

[23]If K^* exists, then it is unique in the range where μ^a is increasing in K.

First note that if GIRI did not hold ($\mu' < 0$), D would be negative and the steady state would be saddle path stable ($\lambda_1\lambda_2 < 0$). But under GIRI, it looses its stability. For low levels of congestion it is totally unstable, and the economy may converge to endogenous fluctuations around a limit cycle (which obviously satisfy the transversality condition (2.3.8)). To see this note that B depends negatively on E_K, which depends positively on the congestion parameter ξ. Assuming there is a range of ξ in which B goes from positive to negative as ξ increases, it can be seen that a Hopf bifurcation occurs at the critical value $\xi*$ where $B = 0$. Whether a stable limit cycle appears on the side $\xi < \xi*$ where congestion is low depends on a stability formula which should be interpreted in specific examples.

In order to gain some intuition about the role of returns on investment in the determination of dynamics, write down the second equation of system (2.3.7) close enough to steady state so that $I \approx 0$ and one may ignore the term for expected changes in investment "costs":

$$\dot{I} \approx \frac{V_i}{V_{ii}}\left(\rho - \mu^a\right).$$

Now note that if $\mu^a{}_K < 0$, μ^a is higher than ρ when $K < K*$, and lower afterwards; but if $\mu^a{}_K > 0$, the reverse is true. Thus under decreasing (adjusted) returns the path of investment will be falling if K is below its steady state value, while under increasing returns it will be rising. This implies that if $\mu^a{}_K < 0$ there must exist a saddle path where agents start with positive investment when $K < K*$, then reduce it to zero as $K*$ is approached. But if $\mu^a{}_K > 0$, agents' behavior becomes in a sense perverse because they will be following a path of *rising* investment when $K < K*$, so that $K*$ will necessarily be reached with positive investment and the economy will overshoot its steady state. Beyond $K*$ investment starts falling and tends to reverse the growth, but by a similar argument the economy will also undershoot $K*$ in the following contraction, and will be involved in repeated cycles thereafter.

2.4. *Multiple Equilibria*

The presence of increasing returns introduces two sources of multiple equilibria in the class of models (2.3.4), and destroys the neoclassical uniqueness result on both accounts: (1) The first reason for multiplicity is the *indeterminacy* that arises from the very nature of dynamics around a stable limit cycle; (2) The second is the *steady state multiplicity* that could result if GIRI only holds locally, i.e., if there are regions of increasing returns and others of decreasing returns. These two sources are examined in turn.

Indeterminacy. The economy's convergence to a *stable* limit cycle makes economic equilibrium indeterminate. As can be seen from figure 2, there is a *continuum* of equilibrium paths that converge to the limit cycle, all of which satisfy the FOC (2.1.8) and the transversality condition (2.3.8). Different equilibria correspond to different states of expectations as to the future course of the economy that are "self-fulfilling": more optimistic forecasts of aggregate investment induce more optimistic forecasts for returns on investment, which in turn induce a higher level of firm investment. This gives expectations a degree of freedom by which they may be determined exogenously. If the exogenous determinant of expectations are stochastic "animal spirits," it gives rise to an additional source of instability in the economy due to changes in the exogenous determinants of expectations. (Woodford (1984) shows how "sunspot" equilibria can be constructed in the presence of indeterminacy[24]). But one should be careful to distinguish this *additional* possible exogenous source of instability from the more fundamental endogenous steady state instability that leads to fluctuations even under perfect foresight.

Multiple Steady states. The other reason for multiple equilibria arises if we can only maintain the GIRI assumption locally. In this case there is no reason to believe that, globally speaking, models of the form (2.3.4) must have a unique steady state. It is quite possible that the "adjusted" return on investment μ^a defined in (2.3.6) not be a monotone function of K, in which case $\mu^a(K, 0)$ may equal the discount rate ρ at more than one steady state level of capital. If μ^a is continuous in K, it will necessarily be alternately increasing and decreasing in K at the different steady state levels. This means that the steady states will typically be alternately unstable (possibly with a limit cycle) and saddle path stable.

In order to explore how this possibility may arise in actual economies consider the first example of the decentralized Ramsey economy in which $\mu^a = \mu$. Further assume that GIRI arises because of the *internal* economies of scale described in section 2.1. It is possible that the conditions for increasing returns on investment may only hold for a certain range of K. This will happen if there is an upper limit to the amount of labor households can supply, which means that there is a full employment level of capital K^f beyond which aggregate employment cannot rise (this feature is included in the appendix). For $K < K^f$, firms can take advantage of increasing returns by adding both capital an labor as they expand, and

[24]See also Howitt and McAfee (1988b) who construct an animal spirits cycle in a special case of the class of models (2.3.4), and prove its stability under learning.

$\mu(K)$ is increasing; but for $K > Kf$, employment can no longer be increased, firms suffer from the decreasing marginal productivity of their capital, and $\mu(K)$ is decreasing.

The increasing-decreasing property of the $\mu(K)$ schedule implies that there can be *two* equilibria where $\mu = \rho$. The first (K^*) occurs when $\mu(K)$ is increasing. It is unstable with outward-spiraling dynamics which possibly converge to a limit cycle. The second (K^{**}), which is higher, occurs when $\mu(K)$ is decreasing, and is saddle path stable. This is illustrated in figure 3.

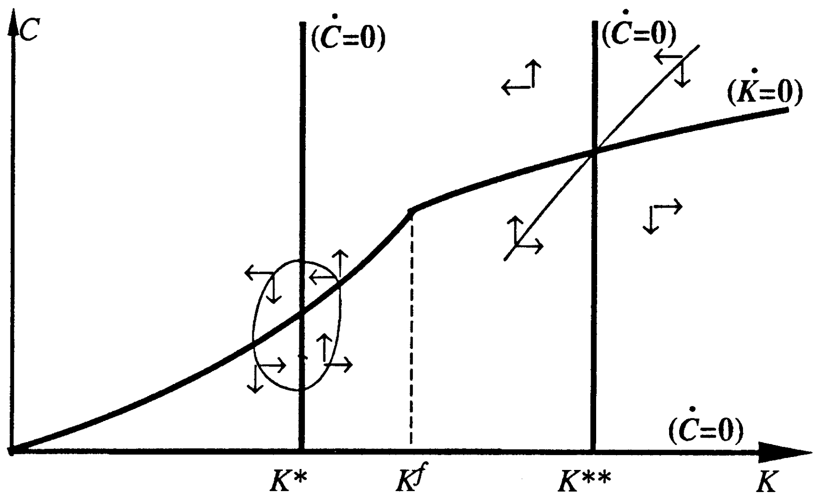


FIGURE 3

A tentative argument[25] can be made that one expects the economy to converge to the limit cycle rather than to the saddle path stable equilibrium. This is because there is a continuum of paths that converge to the limit cycle, but only one saddle-path that converges to K^{**}. Thus there is a sense in which a high degree of coordination is needed for agents to place themselves on the unique saddle path. This suggests that although there exists an equilibrium in which the economy is stable and remains at full employment, because of lack of coordination it is likely to be caught in a state of perpetual cyclical fluctuations in output and employment. The argument is only tentative, however, because the way the economy comes to be on one path rather than another is not modeled explicitly. From this point of view, endogenous fluctuations are a chronic consequence of the economy's tendency to get stuck, because of lack of coordination, in a mode of underemployment with too little capital for the existing labor force.

[25]This argument was presented by Howitt and McAfee (1988a) in a similar context.

3. CONCLUSION

This essay has analyzed the implications of increasing returns on the economy's short-run dynamics. Scale economies were found to have important dynamic implications because they can give rise to global increasing returns on investment (GIRI). These may arise under either external or internal economies. In the first case, GIRI holds as long as external economies are not overcome by internal diseconomies. In the second case, GIRI holds if firms are able to take advantage of increasing returns in the aggregate by increasing both capital and labor as they expand. This was found to require some degree of "real wage rigidity" that keeps labor affordable as the economy expands.

Compared to the saddle path stability result under constant returns, GIRI has a destabilizing effect on the economy's steady state. Thus under GIRI some form of expansion diseconomies is needed to explain the stationarity of short-run dynamics. In their presence the economy will converge to a stable limit cycle around steady state if a certain stability condition is satisfied. Moreover, by the very nature of dynamics around a stable limit cycle, equilibrium paths will be indeterminate and "animal spirits" may have important effects. These results may remain valid even if GIRI only holds locally and returns to scale are ultimately decreasing: Despite the existence of another stable equilibrium path, the economy may still be caught, for lack of coordination, in a state of perpetual fluctuation around an unstable steady state.

In sum, the presence of increasing returns suggests that, beyond the effect of exogenous shocks—monetary, technological, psychological, or other—there may be a more fundamental source of instability in industrial economies rooted in the behavior of optimizing agents and the kind of technology in which they operate.

APPENDIX

In this appendix a detailed formal model similar to Kiyotaki's (1985) is developed that encompasses most of the issues discussed in the main text. The richness of the argument—which involves optimal consumption and investment behavior, increasing returns, monopolistic competition, efficiency wages, congestion externalities—makes for a complex model that is unsuitable for expository purposes and is only presented here as a reference. In the second section are presented a rigorous proof of the existence of a limit cycle in this economy.

A.1. *The Model*

Intertemporal Optimization by Households and Firms: Households in this economy are identical, and their number is normalized to $N = 1$ for simplicity. The infinitely lived representative household (whose behavior stands for both individual and aggregate household behavior) supplies labor inelastically and faces at time t the utility maximization problem

$$\underset{\{C(s), A(s)\}}{Max} \int_t^{\infty} U(C(s))e^{-\rho(s-t)}ds \quad \text{s.t.} \quad \dot{A}(s) = r(s)A(s) + w(s)L(s) - C(s),$$

taking $A(t)$ and the paths of real interest rates $r(s)$, wages $w(s)$, and employment $L(s)$ as given. The dynamic budget constraint equates the increase in the real value of its financial assets A to its financial income rA plus its labor income wL (w is the real wage rate) minus its real consumption level C. The household's employment level can neither exceed its exogenously given labor endowment L^f nor the firms' labor demand L^D:

$$L = Min\{L^f, L^D\}.$$

The application of Pontriagin's maximum principle (see Arrow (1969)) yields the FOC:

$$\frac{\dot{C}}{C} = \sigma(C)(r-\rho), \quad \sigma(C) = -\frac{U'(C)}{U''(C)}\frac{1}{C} > 0. \tag{A.1}$$

Output is produced by $m = 1, ..., M$ (later normalized to 1) symmetric infinitely lived firms with instantaneous maximum profit function $\pi(k; H)$ (to be derived later), which depends on the representative firm's capital stock k and a set H of of economic aggregates. Firms incur instantaneous convex adjustment costs:

$$g(i, \xi K) = \gamma(i) + \xi iK, \quad \gamma', \gamma'' > 0, \quad \xi > 0, \tag{A.2}$$

when they want to raise the capital stock at the net rate $i = dk/dt$. Because of congestion externalities, which are measured by the parameter ξ, these costs depend positively on the aggregate capital stock K. The representative firm's value maximization problem at time t is:

$$\underset{\{i(s)\}}{Max} \int_t^{\infty} [\ \pi(k(s); H(s)) - g(i(s), \xi K(s))\ ] e^{-\int_t^s r(u)du} ds \quad \text{s.t.} \quad \dot{k}(s) = i(s),$$

taking $k(t)$ and the paths of $H(s)$, $r(s)$ as given. Again, assuming a solution exists for this problem, by the maximum principle it must satisfy:

$$\frac{di}{dt} = \frac{g_i}{g_{ii}}\left[r - \frac{1}{g_i}(\mu + \xi I)\right],$$

where $\mu = \partial\pi/\partial k$ is the marginal profitability of capital and I is aggregate investment. When the aggregate counterpart of this FOC is needed, I will normalize the number of firms to $M = 1$ for simplicity, so that:

$$\dot{I} = \frac{g_i}{g_{ii}}\left[r - \frac{1}{g_i}(\mu + \xi I)\right], \tag{A.3}$$

The Demand for Goods: The representative household's level of consumption C is in fact a function of its consumption C_m of the differentiated products of each firm m. I assume a symmetric CES form for this subutility function (as in Dixit and Stigliz (1977)):

$$C = M^{-\beta}\left(\sum_{m=1}^{M} C_m^{\frac{1}{1+\beta}}\right)^{1+\beta}, \quad \beta > 0.$$

The household minimizes the expenditure required to achieve any given level of consumption C, taking goods prices $\{P_m\}$ as given:

$$\underset{\{C_m\}}{Min} \sum_{m=1}^{M} P_m C_m \quad \text{s.t.} \quad C = M^{-\beta}\left(\sum_{m=1}^{M} C_m^{\frac{1}{1+\beta}}\right)^{1+\beta}.$$

Given that firms price symmetrically, we can define the aggregate price level in terms of any good in the economy to be unity:

$$P = P_m = 1, \text{ for all } m.$$

In this case the first order conditions to the above expenditure minimization problem yield the following demand function for firm m's product:

$$C_m = M^{-1}\left(\frac{P_m}{P}\right)^{-\frac{1+\beta}{\beta}} C. \tag{A.4}$$

For simplicity assume that the same goods are used for consumption and capital investment, and that the representative firm's investment index i depends on its investment in specific goods in the same manner as household consumption. If firms take all prices as given for the sake of investment, their investment demand i_m for the particular good m will be as in (A.4):

$$i_m = M^{-1}\left(\frac{P_m}{P}\right)^{-\frac{1+\beta}{\beta}} i. \tag{A.5}$$

Now using (A.4) and (A.5) we can express the demand for each product y_m in terms of the economic aggregates C and $I = Mi$:

$$y_m = M^{-1}\left(\frac{P_m}{P}\right)^{-\frac{1+\beta}{\beta}} (C + I). \tag{A.6}$$

Note that by aggregating (A.6) over all firms, we get the usual aggregate goods market equilibrium condition:

$$Y = C + I, \tag{A.7}$$

where aggregate output is given under symmetry ($y_m = y$, for all m) by $Y = My$.

The Supply of Goods and the Labor Market: Of the many ways mentioned in section 1.2 of generating underemployment and a fixed real wage, I choose the efficiency wage story mainly because it leads to a simple analysis of dynamics. (Under the voluntary unemployment story, the level of employment enters the instantaneous utility function $U(C-w^*L)$ and we would have to take the dynamic behavior of employment into account. Although this would make the analysis of dynamics more complex, it would not affect any of the main results). For this purpose assume that the representative firm's wage rate w enters its production function $F(k, l)$ in a labor augmenting way (as in Solow (1979)):

$$y = F(k, e(w)l^D)$$

where l^D is the amount of labor hired by the firm, and $e(w)$ is the workers' level of effort, which increases with the real wage rate. The elasticity $\eta(w) = (de/dw)(e/w)$ of the level of effort with respect to the wage rate is assumed to be a decreasing function of w that goes to zero as w increases.

Facing this production function, the demand function (A.6), its capital stock k_m, and the aggregates $H = \{Y, P, w\}$, each firm m maximizes its instantaneous profits :

$$(P_m/P)F(k_m, e(w)l_m^D) - w_m l_m^D$$

to obtain the optimal profit function $\pi(k_m; H)$ which was referred to earlier.

Under these circumstances one can show that the equilibrium wage rate will be equal to the "efficiency" wage w^* which maximizes worker effort per dollar $e(w)/w$ (i.e., $\eta(w^*) =$

1) as long as firms can hire all the labor they want at w^*. Otherwise, if labor supply falls short of labor demand at w^*, then the wage rate must rise above w^* ($\xi < 1$) to clear the labor market. Thus under symmetry the marginal productivity of labor, which by the first order condition must be set equal to:

$$F_l(k, e(w)l^D) = (1+\beta)w/e(w), \tag{A.8}$$

determines the firm's employment level l^D (with $w = w^*$) as long as full employment has not been reached ($l^D < L^f/M$), and determines the wage rate that clears the labor market under full employment ($l^D = L^f/M$).

Finally, by applying the envelope theorem to the above maximization problem, we find that under symmetry the marginal profitability of capital is given by:

$$\mu = (1+\beta)^{-1}F_k(k, e(w)l^D). \tag{A.9}$$

Determination of Economic Aggregates: This economy's dynamics are two dimensional and can be studied in (K, I)-space. To do so I now turn to the way K and I determine the economic aggregates L, μ, Y, C, and r. I first need to introduce the representative production function. Assume it has the functional form:

$$F(k, l; Y) = qk^a l^b - vY, \quad q, v > 0, \quad 0<a<1, \quad 0<b<1, \quad 1<a+b<(1+\beta). \tag{A.10}$$

Under this specification $a+b > 1$ is equivalent to assuming increasing marginal productivities with scale. The upper limit $(1+\beta)$ on $(a+b)$ guarantees that the SOC to profit maximization are satisfied. In the same way I introduced a congestion effect on adjustment costs, I assume that congestion externalities affect firm output by assuming that it depends negatively on aggregate output. The "congestion" parameter v measures the strength of this effect.

Using this production function and equation (A.8) we can calculate employment as a function of the aggregate capital stock:

$$L(K) = \begin{cases} \theta_L K^{\alpha}, \text{ if } K \le K^f; \\ L^f, \text{ otherwise,} \end{cases} \quad \text{where } \alpha = \frac{a}{1-b}; \ \theta_L = \left(\frac{bqe(w^*)^b}{M^{a+b-1}(1+\beta)w^*} \right)^{\frac{1}{1-b}}; \ K^f = \left(\frac{L^f}{\theta_L} \right)^{\frac{1}{\alpha}}.$$

Employment increases with the level of capital until full employment is reached at K^f, after which the economy remains at full employment.

From this last equation and (A.9) we obtain the marginal profitability of capital as a function of the aggregate capital stock:

$$\mu(K) = \begin{cases} \theta_\mu K^{\alpha-1}, \text{ if } K \le K^f; \\ \theta^f_\mu K^{a-1}, \text{ otherwise,} \end{cases} \quad \text{where } \theta_\mu = \frac{a}{b} w^* \theta_L; \ \theta^f_\mu(e) = \left[\frac{e(w)}{e(w^*)} \right]^b \theta_\mu \left(K^f \right)^{\frac{b}{1-b}}.$$

As long as full employment has not been reached, the economy takes advantage of increasing returns and μ is an increasing function of the capital stock ($\alpha > 1$ is equivalent to $a+b > 1$). Once the full employment capital stock K^f is reached, $\mu(K)$ becomes a decreasing function because of the decreasing marginal productivity of capital ($a < 1$).

Aggregate output can be obtained by inserting $L(K)$ into (A.10):

$$Y(K;\,\phi) = \begin{cases} \phi\theta_Y K^{\alpha}, \; if\, K \le K^f; \\ \phi\theta_Y^f K^{a}, \; \text{otherwise}, \end{cases} \quad \text{where} \quad \phi = \frac{1}{1+Mv}; \quad \begin{matrix} \theta_Y = \dfrac{(1+\beta)}{a}\theta_\mu; \\ \theta_Y^f(e) = \dfrac{(1+\beta)}{a}\theta_\mu^f(e). \end{matrix} \tag{A.11}$$

It is a convex function of the capital stock ($\alpha > 1$) as long as there remain unemployed labor resources for the economy to draw from and thus take advantage of increasing returns. When full employment is reached at K^f no additional labor can be added and the economy suffers from decreasing marginal returns to capital ($a < 1$). $Y(K)$ becomes concave. ϕ is an inverse measure of congestion: as the degree of congestion v increases from 0 to ∞, ϕ decreases from 1 to 0. Thus Y and Y' fall with congestion.

The determination of output together with the goods market equilibrium condition give us immediately aggregate consumption as a function of (K, I):

$$C(I, K) = Y(K) - I.$$

Finally I turn to the determination of the interest rate r. For this purpose, differentiate the goods market equilibrium equation (A.6) with respect to time (noting that $dK/dt = I$):

$$Y'(K)I = dC/dt + dI/dt.$$

If we think of this equation together with the household and firm FOCs (A.1) and (A.3) as three equations in the three unknowns r, dC/dt, and dI/dt, we can solve for r:

$$r(K, I) = \frac{1}{\kappa_U + \kappa_g}\left[Y'I + \kappa_U\rho + \kappa_g\frac{(\mu + \xi I)}{g_i}\right], \tag{A.12}$$

where κ_U and κ_g are the *inverted* curvatures of the utility and adjustment cost functions:

$$\kappa_U = -\frac{U'}{U''} \quad \text{and} \quad \kappa_g = \frac{g_i}{g_{ii}}.$$

The first term $Y'I$ in the expression for r represents changes in aggregate income ($dY/dt = Y'I$). If this term were zero, the interest rate would be a weighted average of the households' discount rate ρ and firms' return on investment $(\mu + \xi I)/g_i$, adjusted for present and foreseen changes in adjustment costs:

$$r = \left[\theta\rho + (1-\theta)\left(\frac{\mu + \xi I}{g_i}\right)\right], \quad \text{where} \quad \theta = \frac{\kappa_U}{\kappa_U + \kappa_g} \in [0, 1]. \tag{A.13}$$

The weights are heavier, the smaller is the curvature of the corresponding agent's maximization problem. By the FOCs (A.1) and (A.3), the fact that r is between ρ and $(\mu + \xi I)/g_i$ implies that one of C and I will be increasing and the other decreasing so that $dC/dt + dI/dt = dY/dt = 0$. But if dY/dt is not zero this need not hold, and r can be below or above both ρ and $(\mu + \xi I)/g_i$. If income is rising, for example, then it is possible for both C and I to be rising simultaneously, and thus for r to be above both ρ and $(\mu + \xi I)/g_i$.

Equilibrium Dynamics: We are now ready to analyze the system's dynamics in (K, I)-space, which are governed by the capital accumulation equation and the FOC (A.3):

$$\dot{K} = I;$$
$$\dot{I} = \frac{g_i(I, \xi K)}{g_{ii}(I, \xi K)}\left[r(K, I) - \frac{(\mu(K) + \xi I)}{g_i(I, \xi K)} \right]. \qquad \text{(A.14)}$$

Despite the fact that the functions $r(K, I)$ and $\mu(K)$ are not continuously differentiable, we are guaranteed the existence and uniqueness of solutions because the system is Lipschitzian (see Sansone and Conti (1964), pp. 13-15).

I we define the return on investment adjusted for present installation costs as in (2.2.6):

$$\mu^a(K, I; \xi) \equiv \frac{\mu(K)}{g_i(I, \xi K)},$$

then a steady state other than the origin[26] must satisfy:

$$I = 0 \quad \text{and} \quad \mu^a(K, 0; \xi) = \rho.$$

From here on, I assume whenever I need to that the "adjusted" return on investment $\mu^a(K, 0; \xi)$ depends on K in the same way a $\mu(K)$ (see section 2.2 for a discussion of this assumption). In this case, as long as L_f is large enough, it can be seen that there are two levels of K that satisfy the steady state conditions. One $K^* < K_f$ is an underemployment equilibrium that occurs when $\mu' > 0$, the other $K^{**} > K_f$ has full employment and occurs when $\mu' < 0$. The phase portrait is given in figure 3 of the main text.

The eigenvalues at any of these steady states are given by:

$$\lambda = \frac{1}{2}\left(B \pm \sqrt{B^2 - 4D}\right),$$
$$\text{where } B = (1-\theta)Y' + \theta\left(\rho - \kappa_g \frac{\xi}{g_i}\right) \text{ and } D = \kappa_g \theta \mu^a_K \qquad \text{(A.15)}$$

These correspond to the eigenvalues (2.1.10) derived in the first example when adjustment costs are linear $g(i, \xi K) = i$—i.e., when $\kappa_g \to \infty$ (and thus $\theta \to 0$), $g_i = 1$, and $\xi = 0$. They correspond to the eigenvalues (2.2.8) of the second example (section 2.2) when

---

[26]The origin is also a steady state if $\sigma(0)$ is finite.

instantaneous utility is linear—i.e., when $\kappa_U \to \infty$ (and thus $\theta \to 1$). Since $\mu^a{}_K$ is assumed to have the same sign as μ', D is positive for K^* and negative for K^{**}. Since for "reasonable" (low) congestion levels $B > 0$, this implies that K^* is totally unstable and K^{**} is saddle path stable. The next section derives conditions under which a limit cycle will appear around K^*.

A.2. *Existence of a Limit Cycle*

The proof of existence of a limit cycle around K^* in system (A.14) relies on the Hopf bifurcation theorem (see Guckenheimer and Holmes (1983), pp. 151-2):

Theorem (Hopf): *Suppose that the system $dz/dt = h_\psi(z)$, $z \in \Re^n$, $\psi \in \Re$, has an equilibrium (z^*, ψ^*) at which the following properties are satisfied:*

(i) $D_z f_{\psi^*}(z^*)$ has a simple pair of pure imaginary eigenvalues and no other eigenvalue with zero real parts. This implies that there is a smooth curve of equilibria $(z(\psi), \psi)$ with $z(\psi^*) = z^*$. The eigenvalues $\lambda(\psi)$, $\overline{\lambda}(\psi)$ of $D_z h_{\psi^*}(z(\psi))$ which are imaginary at $\psi = \psi^*$ vary smoothly with ψ. If moreover:

(ii) $d\mathrm{Re}\lambda(\psi)/d\psi\,|_{\psi=\psi^*} > 0$,

then there exists a periodic solution bifurcating from z^* at $\psi = \psi^*$ and the period of the solutions is close to $2\pi / |\lambda(\psi^*)|$.

The eigenvalues of (A.14) at $z^* = (K^*, 0)$ are given by (A.15). Since $D > 0$ by assumption, a bifurcation occurs when $B = 0$. Since $0 \leq \theta \leq 1$, this will happen if

$$Y'(K^*;\phi) = 0 \quad \text{and} \quad \rho = \xi/g_{ii}(0, \xi K^*).$$

In the first example ($\theta = 0$) only the first condition matters and if we choose the bifurcation parameter to be $\psi = \phi$, a Hopf bifurcation occurs by (A.11) at the critical value $\phi^* = 0$. In the second example ($\theta = 1$) it is the second condition that matters. If we choose $\psi = \xi^{-1}$ then by (A.2) a bifurcation occurs at the critical value corresponding to $\xi^* = \rho/\gamma'(0)$. Condition (ii) of the theorem is satisfied in both cases. Finally in the intermediate case where $0 < \theta < 1$, if we think of $\phi(\psi)$ an $\xi(\psi)$ as depending (respectively positively and negatively) on a third parameter ψ in such a way that they reach their critical values ϕ^* and ξ^* simultaneously when $\psi = \psi^*$, then a bifurcation occurs at ψ^*. The theorem also gives us an approximation for the period of the cycle near ψ^*, which is equal to $2\pi/|\lambda(\psi^*)|$ where from (A.15):

$$|\lambda(\psi^*)| = \sqrt{\kappa_g \theta \mu_K^a}.$$

The limit cycle will exist in a neighborhood on one side of the critical value. The realistic case is when it exists for $\psi > \psi^*$(low congestion). To address this issue, the following result can be used (see Guckenheimer and Holmes (1983), p. 152-3; or Marsden and McCracken (1976), pp. 131-5):

Proposition: *Suppose the system $dz/dt = h_\psi(z)$ described in Hopf's bifurcation theorem is two-dimensional and that, possibly after a change of coordinates, it can be written in the form:*

$$\dot{z} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = h(z) \quad \text{with} \quad D_z h(z^*) = \begin{bmatrix} 0 & |\lambda^*| \\ -|\lambda^*| & 0 \end{bmatrix},$$

where $|\lambda^*| = |\lambda(\psi^*)|$. Also let the stability coefficient "a" be defined as:

$$a = \left[f_{xxx} + f_{xyy} + g_{xxy} + g_{yyy}\right] - \frac{1}{|\lambda^*|}\left[f_{xy}(f_{xx} + f_{yy}) - g_{xy}(g_{xx} + g_{yy}) - f_{xx}g_{xx} + f_{yy}g_{yy}\right].$$

Then if $a < 0$, the periodic solutions appear when $\psi > \psi^*$ and are stable (attracting) limit cycles. If $a > 0$, they appear when $\psi < \psi^*$ and are unstable (repelling).

I will derive the stability condition $a < 0$ in the two limit cases when $\theta = 0$ and 1. When $\theta = 0$ the dynamics are given by system (2.1.9) of the first example. The change of coordinates:

$$x = K, \quad y = -C/|\lambda^*|$$

gives us the functions

$$\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} Y(x) + |\lambda^*|y \\ \sigma y(\mu(x) - \rho) \end{bmatrix}$$

which satisfy the condition in the proposition. Under these new coordinates we get:

$$a = \left(Y''' + \sigma\mu''\right) - \sqrt{\sigma C^* \mu'}\left(\sigma\mu'' + \frac{Y''\mu''}{\mu'}\right).$$

By (A.11) $Y'' = Y''' = 0$ at $\psi = \psi^*$ and the condition $a < 0$ is therefore:

$$\mu''\left(1 - \sqrt{\sigma C^* \mu'}\right) < 0.$$

It is discussed in section 2.1.

In the case when $\theta = 1$, the dynamics are governed by system (2.2.7) of the second example. The coordinate change:

$$x = -|\lambda^*|K, \quad y = I$$

allows us to write the system in a form that satisfies the condition in the proposition. Using (A.2):

$$\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} -|\lambda^*|y \\ \frac{1}{\gamma''(y)}\left(\rho\left(\gamma'(y) - \xi^* x/|\lambda^*|\right) - \left(\mu\left(-x/|\lambda^*|\right) + \xi^* y\right)\right) \end{bmatrix}.$$

The stability coefficient is given by:

$$a = \rho \frac{\gamma'''}{\gamma''}\left(1 - 2\frac{\gamma'''}{\gamma''}\right),$$

and the condition $a < 0$ is therefore as given in (2.2.12).

REFERENCES

ARROW, K. J. (1969): "Applications of Control Theory to Economic Growth," in *Mathematics of the Decision Sciences, Part 2,* ed. by G. B. Dantzig and A. F. Veinott. Providence, Rhode Island: American Mathematical Society.

AZARIADIS, C. (1981): "Self-fulfilling Prophecies," *Journal of Economic Theory,* 25, 380-396.

BARRO, R., AND R. G. KING (1984): "Time-Separable Preferences and Intertemporal-Substitution Models of the Business Cycle," *Quarterly Journal of Economics,* 99, 817-39.

BILS, M. (1985): "Real Wages over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy,* 93, 666-89.

DIAMOND, P. (1982): "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy,* 90, 881-894.

DIAMOND, P. AND D. FUDENBERG (1987): "Rational Expectations Business Cycles in Search Equilibrium," *Journal of Political Economy*, forthcoming.

DIXIT, A. AND J. E. STIGLIZ (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67, 297-308.

EISNER, R., AND R. H. STROTZ (1963): "Determinants of Business Investment," in Commission on Money and Credit, *Impacts of Monetary Policy*. Englewood Cliffs, New Jersey: Prentice-Hall, pp. 59-337.

FRISCH, R. (1933): "Propagation Problems and Impulse Problems in Dynamic Economics," in *Economic Essays in Honor of Gustav Cassel*. London: George Allen and Unwin.

GABISCH, G., AND H.-W. LORENZ (1987): *Business Cycle Theory*. New York-Berlin-Heidelberg: Springer-Verlag.

GEARY, P., AND J. KENNAN (1982): "The Employment-Real Wage Relationship: An International Study," *Journal of Political Economy,* 90, 850-871.

GOODWIN, R. M. (1951): "The Non-Linear Accelerator and the Persistence of Business Cycles," *Econometrica*, 19, 1-17.

GOULD, J. P. (1968): ""Adjustment Costs in the Theory of Investment of the Firm," *Review of Economic Studies*, 35, 47-55.

GUCKENHEIMER, J., AND P. HOLMES (1983): *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York-Berlin-Heidelberg: Springer-Verlag.

HALL, R. E. (1988): "Increasing Returns: Theory and Measurement with Industry Data," mimeo, Stanford University.

HICKS, J. R. (1950): *A Contribution to the Theory of the Trade Cycle*. Oxford: Clarendon Press.

HIRSCH, M. W., AND S. SMALE (1974): *Differential Equations, Dynamical Systems and Linear Algebra*. New York: Academic Press.

HOWITT, P. (1985): "Transaction Costs in the Theory of Unemployment," *American Economic Review,* 75, 88-100.

——— AND R. P. McAFEE (1988a): "Stability of Equilibria with Trade-Externalities," *Quarterly Journal of Economics,* 103, 261-277.

——— AND R. P. McAFEE (1988b): "Animal Spirits," mimeo, University of Western Ontario.

KALDOR, N. (1972): "The Irrelevance of Equilibrium Economics," *Economic Journal,* 82, 1237-1255.

KATZ, L. F. (1987): "Efficiency Wage Theories: A Partial Evaluation," *NBER Macroeconomics Annual*, 1, 235-276.

KIYOTAKI, N. (1985): "Multiple Expectational Equilibria under Monopolistic Competition," *Quarterly Journal of Economics*, 103, 695-741.

KYDLAND, F., AND E. C. PRESCOTT (1982): "Time to Build and Aggregate Fluctuations," *Econometrica,* 50, 1345-1370.

LUCAS, R. E., JR. (1967a): "Adjustment Costs and the Theory of Supply," *Journal of Political Economy,* 75, 321-334.

——— (1967b): "Optimal Investment Policy and the Flexible Accelerator," *International Economic Review,* 8, 78-85.

MARSDEN, J. E., AND M. McCRACKEN (1976): *The Hopf Bifurcation and Its Applications*. New York-Berlin-Heidelberg: Springer-Verlag.

PRATTEN, C. F. (1971): *Economies of Scale in Manufacturing Industry*. London-New York: Cambridge University Press.

RAMSEY, F. P. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 88.

ROMER, P. (1986): "Increasing Returns and Long Term Growth," *Journal of Political Economy,* 94, 1002-1037.

——— (1987): "Growth Based on Increasing Returns Due to Specialization," *American Economic Review,* 77-2, 56-62.

SANSONE, G., AND R. CONTI (1964): *Nonlinear Differential Equations*. Oxford-New York: Pergamon Press.

SLUTZKY, E. (1937): "The Summation of Random Causes as the Source of Cyclic Processes," *Econometrica*, 5, 105-146.

SOLOW, R. M. (1979): "Another Possible Source of Wage Stickiness," *Journal of Macroeconomics*, 1, 79-82.

TREADWAY, A. B. (1969): "On Rational Entrepreneurial Behavior and the Demand for Investment," *Review of Economic Studies,* 36, 227-239.

UZAWA, H. (1965): "Optimum Technical Change in an Aggregative Model of Economic Growth," *International Economic Review*, 6, 18-31.

WOODFORD, M. (1984): "Indeterminacy of Equilibrium in the Overlapping Generations Model: A Survey," mimeo, Columbia University.

YOUNG, A. A. (1928): "Increasing Returns and Economic Progress," *Economic Journal,* 38, 527-542.

ZARNOWITZ, V. (1985): "Recent Work on Business Cycles in Historical Perspective," *Journal of Economic Literature*, 23, 523-580.

ESSAY 2

INDETERMINACY AND INSTABILITY IN MACRO MODELS WITH EXTERNAL EFFECTS

Macroeconomists have recently focused their attention on different reasons why aggregate quantities can affect agents' payoffs in ways that are not fully reflected in prices, and on the possibility that these "external effects" can generate Keynesian-style market failures. Models of search which produce "thick market" externalities (Diamond (1982), Howitt (1985)), or of imperfect competition where agents look at quantity as well as price signals (Heller (1986)), have been shown to generate multiple equilibria with high and low levels of activity, but no individual incentive to move unilaterally to the more desirable equilibrium. Although these models run entirely in "real" terms, they can be seen as important complements (rather that contenders) to more traditional Keynesian models based on nominal rigidities, especially since the kind of external effects they rely upon can generate potentially large effects of aggregate demand on output.

In an effort to understand the kind of mechanism driving these results, Cooper and John (1988) have characterized the kind of external effects that are needed by what they call "strategic complementarities": an aggregate increase in activity should have a positive effect on agents' incentive to produce. In this manner, optimism about a high activity level can raise individuals' incentives and become self-fulfilling, and pessimism can lower incentives and lead to depression. This essay is an attempt to further our insight into this phenomenon and the kind of economies where it may appear, and to draw on more recent research to explore its dynamic implications.

The first part derives a stronger necessary condition for external effects to give rise to Keynesian market failures. Indeed, as Cooper and John emphasize, the mere presence of strategic complementarities is far from a guarantee for multiple equilibria. Even though widespread imperfect competition, for example, can give rise to aggregate demand spillovers (and therefore strategic complementarities in output), only in special cases will this lead to the possibility of coordination failures. What is needed, I argue, is for the positive external effect of expansion on incentives to be strong enough so as to overcome any eventual negative internal effect. In this way, a simultaneous increase in economic activity will leave all agents with a higher incentive to produce. In other words, the individual marginal return on effort will rise as agents in the economy expand their output simultaneously.

A systematic review is given of the different ways in which this *social increasing returns* condition could arise. Increasing returns in individual production are clearly a possible source (Kiyotaki (1988), Murphy, Shleifer, and Vishny (1988a)), but they are not the only one. Other possibilities are external economies (e.g. because of "thick market" externalities), or even countercyclical markups (since smaller markups typically mean more incentive to produce).

The second part of this essay addresses the question of the dynamic implications of multiple equilibria in the context of an abstract representative agent economy. Again, the key to this question is that the economy must exhibit social increasing returns around at least one of the equilibria. This has a *destabilizing* effect on dynamics because it implies that agents have an incentive to increase individual output more than proportionally in response to an increase in aggregate activity (a proportional response leaves them with higher incentives and willing to produce more).

One reaction to this problem is relief: the economy cannot be observed around an unstable equilibrium, and must therefore converge to a neighboring stable equilibrium. But the possibility of fluctuations in the unstable region cannot be so easily dismissed. In the presence of natural limits to expansion (capacity constraints, expansion diseconomies, etc.), a plethora of fluctuating equilibrium paths are feasible that remain in the increasing returns region. Moreover, indeterminacy becomes more chronic: in addition to multiple steady states, there can be a continuum of equilibrium paths around them. Examples are presented of the different contexts in which this type of behavior has been studied (Diamond and Fudenberg (1987), Drazen (1988), Howitt and McAfee (1988b), Murphy, Shleifer, and Vishny (1988b), Hammour (1988)). If the external effects in question turn out to be empirically relevant, they have the potential of providing a partial *explanation* for observed macro instability.

The first part of this essay analyzes multiple equilibria in a static model. It derives the necessary condition of social increasing returns, and gives several examples. Part 2 turns to dynamics and discusses the destabilizing effect of increasing returns. Examples are again discussed in detail. Part 3 addresses the question of the possibility of economic fluctuations in the unstable region, if there are limits to expansion. The main results are summarized in the conclusion.

1. MULTIPLICITY

1.1. *A Static Economy*

Consider an economy with $i = 1, \ldots, N$ identical agents, each of whom can exert a level e_i of effort to obtain a net payoff V. Payoffs are a twice differentiable function

$$V = V(e_i, E), \quad V_e > 0, V_{ee} < 0$$

of the individual level of effort e_i, and the average level of effort E in the economy:

$$E \equiv \frac{1}{N}\sum_{j=1}^{N} e_j$$

In this abstract setup, the function V supposedly incorporates the outcomes of all decentralized markets. Although it is possible that the aggregate E affect payoffs V through the operation of some relative price, the interesting results arise when this effect is "external" and comes from some violation of perfect competition, as the examples in section 1.3 make clear. For this reason, the effect of E on V will sometimes simply be referred to as the "external effect."

A representative agent chooses the level of effort e that maximizes his payoff $V(e, E)$, taking the aggregate E as given[1] (assuming N is large, I neglect the effect of his individual level of effort on E). The first order condition for this problem sets to zero the marginal return on effort:

$$V_e(e, E) = 0. \tag{1.1.1}$$

In equilibrium, the individual effort level e^* must satisfy (1.1.1) with $E^* = e^*$.

This equilibrium concept is similar to the Nash concept used by Cooper and John (1988) and and which determines their terminology. But the latter is avoided here because its seems appropriate in a macroeconomic context to assume that agents are too small to take into account the effect of their actions on aggregates. From this perspective, there is no room at a macro level for truly *strategic* interaction as understood by modern dynamic game theory, where, for example, the action of a single agent can trigger a change in behavior by all others (as in Roberts' (1987) macro model, e.g.). This will be particularly relevant to the extension of our equilibrium concept to a dynamic context.

1.2. *Multiple Equilibria*

In order to investigate the conditions under which the external effect in this economy can generate multiple equilibria, define the "reaction function" $e_i(E)$ that gives agent i's optimal level of effort as a function of the aggregate level E. By the concavity of V in e_i, this function is well defined and implicitly given by the first order condition (1.1.1). Thus the slope of the reaction curve is:

$$\frac{de_i}{dE} = -\frac{V_{eE}}{V_{ee}}. \tag{1.2.1}$$

[1] A remark concerning notation: Throughout this paper, lower-case letters (e_i, for example) denote individual vatiables, and upper-case letters (E) denote their economy-wide averages. The index i is sometimes dropped from an individual variable (e) when it refers to a "representative agent."

An equilibrium effort level E^* is one at which $e_i(E)$ crosses the 45° line: $e_i(E^*) = E^*$. First consider the standard case with no external effects ($V_E = 0$). In this case the reaction function is flat, and there can only be one equilibrium (a consequence of the concavity of V). What is needed for multiple equilibria is, at least locally, an upward sloping reaction function, as illustrated in figure 1. This requires that $V_{eE} > 0$ in this region, a condition Cooper and John (1988) term *strategic complementarity* in agents' payoff functions: an increase in the aggregate level of effort E should increase an individual's incentive to produce V_e (which is different from a positive effect on agents' payoffs V)[1].

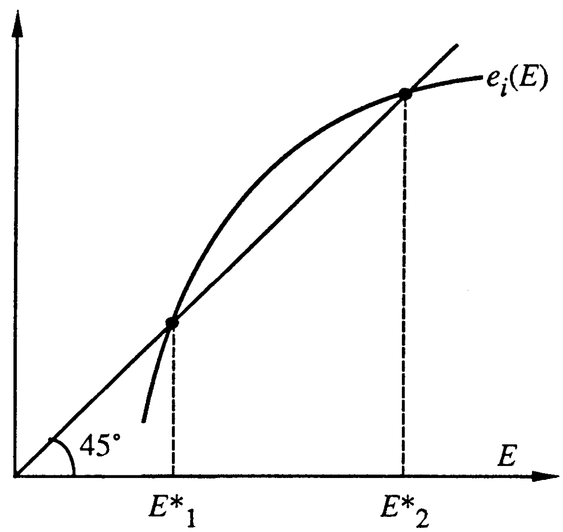


FIGURE 1

In fact, one can derive a stronger necessary condition for multiple equilibria that will prove very useful in determining the kind of economies in which they may appear and will provide a key to understanding their dynamic implications. This condition is based on the simple observation that if there are more that one equilibrium, the reaction curve must cut the 45° line *from below* at least at one of them (see figure 1). This means that its slope must be greater than one, which by (1.2.1) is equivalent to:

$$V_{ee} + V_{eE} > 0. \qquad (1.2.2)$$

Condition (1.2.2) is one about the strength of strategic complementarities, which must be strong enough so that $V_{eE} > -V_{ee} > 0$. But to uncover its economic content, define $R(E)$ as agents' marginal return on effort, when they all exert the effort level $e = E$:

$$R(E) \equiv V_e(E, E).$$

[1]If, moreover, an increase in aggregate activity improves agents' payoffs (i.e. if $V_E > 0$) then equilibria with higher levels of activity dominate those with lower levels.

It is easy to see that (1.2.2) is equivalent to $R'(E) > 0$. In other words, if all agents increase effort simultaneously, the marginal return on their effort should rise. Strategic complementarities must be strong enough so that the economy may exhibit *social (or aggregate) increasing returns* on effort, in the sense that, although the return on effort is decreasing with each individual's effort level ($V_{ee} < 0$, otherwise his maximization problem would not be well-behaved), it is increasing with the aggregate level for the economy as a whole.

The reason why this condition is equivalent to the reaction curve cutting the 45° line from below is worth keeping in mind in view of the forthcoming discussion of dynamics. Start from a symmetric equilibrium E^* where the reaction curve crosses the 45° line, and suppose aggregate effort rises by dE. Then, if the return on effort is increasing in the aggregate ($R'(E^*) > 0$), agent i would react by increasing his effort level *more* than proportionally by $de_i > dE$, because if he were to increase it just proportionally he would end up with a marginal return on effort greater than zero:

$$V_e(E^*+dE, E^*+dE) = R(E^*+dE) \; > \; R(E^*) = V_e(E^*, E^*) = 0,$$

which violates his first order condition (1.1.1). This means that $de_i/dE > 1$, and the reaction curve cuts the 45° line from below. From this, it is clear why social increasing returns will have a destabilizing effect on dynamics. Conversely, if the return on effort is decreasing in the aggregate ($R'(E^*) < 0$), agent i would react by increasing his effort level *less* than proportionally, and the reaction curve cuts the 45° line from above with the resulting stability.

Naturally, increasing returns are only required for a certain range of activity levels. If an economy exhibits multiple equilibria, these will correspond to the reaction curve crossing the 45° line alternatingly from below and from above, and will exhibit alternatingly increasing and decreasing returns. Because of this, if more than one equilibrium is relevant for the economy, then increasing returns must hold *somewhere* in the relevant range of activity. In the extreme case where the economy exhibits increasing returns for all relevant activity levels, equilibrium is again unique, but dynamics will nevertheless be very different from the standard no-externality case.

1.3. *Examples*

From the so far abstract treatment of multiplicity, I now turn to some examples. Blanchard and Summers (1988) usefully classify macro models with multiple equilibria in terms of the intersection of the "labor demand" and "supply" schedules (with a slight abuse of language, these designate the equilibrium wage-employment relationships determined on the demand and supply side). If there are multiple equilibria one schedule must cross the other alternatively from above and from below. This requires a departure from the standard presumptions about these schedules' slope: either labor demand must be (locally) sloping upwards, or supply downwards. In this section I focus on the former case[1] and give a systematic discussion, in light of the increasing returns condition of section 1.2, of three major reasons why labor demand may slope upwards[2]: (1) increasing returns in production; (2) external "thick-market" externalities; and (3) countercyclical markups. These are incorporated in the following setup, based on Kiyotaki's (1985) approach to general equilibrium with imperfect competition.

There are $i = 1, \ldots, N$ symmetric firms, each producing a differentiated good i. The number of households is normalized to $M = 1$, for simplicity. The "representative household" consumes c_i units of each good, supplies l^H units of labor, and derives utility:

$$U(c_1, c_2, \ldots, c_N) - vl^H, \quad v, U_i > 0, U_{ii} \leq 0, \forall i. \tag{1.3.1}$$

A representative firm hires l units of labor to produce $y = f(l)$ units of its good, and derives profits:

$$p(1-\tau)y - wl,$$

where p and w are its price and wage rate (both in utility units), and $\tau = \tau(Y)$ is a transaction cost per unit of output sold that may depend negatively on average output Y in the economy (more on this to come).

In symmetric equilibrium, an equal amount y of each good is produced and average output is $Y = y$. A representative firm takes aggregate Y as given, and faces an inverse demand function $p = p(y, Y)$ that gives its price as a function of own and aggregate output Y. As a monopolist in its own market, it is well known that it will set its price at a markup $1+\mu = 1/[1-1/\varepsilon]$ over its marginal cost (assuming the SOC is also satisfied):

$$p = (1+\mu)\frac{w}{(1-\tau)f_l}, \tag{1.3.2}$$

where $\varepsilon = \varepsilon(y, Y)$ is its demand elasticity.

---

[1] The latter case of downward sloping supply curve could happen within the context of the model in this section if union power that determines real wages gets weaker during booms, a theoretical possibility that was pointed out to me by Robert Solow.

[2] A fourth possible reason discussed by Blanchard and Summers (1988) is "fiscal increasing returns." It is due to the fiscal behavior of a government that increases tax rates in times of high unemployment (and reduced tax base) in order to satisfy fixed spending needs. This mechanism can very easily be incorporated in this section's model by reinterpreting the transaction cost $\tau(Y)$ as a "tax rate" that depends negatively on aggregate income Y.

It will convenient to interpret this markup equation as setting marginal profits to zero and write it in the following form, showing explicitly the dependance on own and aggregate variables:

$$\frac{(1-\tau(Y))f_l(l)}{(1+\mu(y,Y))} - \frac{w}{p(y,Y)} = 0. \tag{1.3.3}$$

In terms of the discussion in the previous two sections, it is clear that (1.3.3) corresponds to the first order condition (1.1.1) that sets the marginal return on "effort" l to zero. The social increasing returns condition for multiple equilibria is therefore that the left-hand side of (1.3.3) must increase locally with a symmetric rise in all firms' employment.

There are three independent mechanisms by which this can happen: (1) if transaction costs τ depend negatively on aggregate output ("thick-market" externalities); (2) if the marginal product of labor f_l increases with employment (increasing returns in production); and (3) if the markup $(1+\mu)$ falls during expansions (countercyclical markups). From a labor market perspective, these effects make "labor demand" upward sloping. The remaining term in (1.3.3) that has not been considered is the real wage w/p and corresponds to labor supply. It would have to *fall* during expansions in order to generate multiplicity, but this is impossible in our model. The best case for multiplicity here is a flat labor supply schedule, when constant real wages do not tend to lower production incentives during booms. Since the maximization of utility (1.3.1) gives the equilibrium real wage $w/p = v/U_i$, (for any $i = 1, \ldots, N$) under symmetry, this corresponds to a constant $U_i(c, c, \ldots, c)$ for all c[1].

External Economies. The first possible source of multiple equilibria are external economies in production or sales (see Weil (1988)). At the risk of trivializing the concept, I have relied in this model on Howitt's (1985) formulation of Diamond's (1982) thick-market externalities in terms of transactions costs. If we assume transaction costs $\tau(Y)$ depend negatively on the level of market activity in the economy ($d\tau/dY < 0$), an increase in aggregate output will reduce transaction costs and tend to increase firms' incentive to produce as measured by the marginal profitability of employment in (1.3.3). Unlike the next two effects, this one does not rely on imperfect competition.

[1]For any sub-utility function $U(c_1, \ldots, c_N)$ define $G(c) \equiv U(c, c, \ldots, c)$. Then the monotonic transformation $G^{-1}[U(c1, \ldots, c_N)]$ (which is not innocent given the additional term for the disutility of work) yields constant real wages without affecting the elasticity of substitution between goods.

Increasing Returns in Production. The second possibility is a convex production function. This can be modeled either in the form of a discrete opportunity to incur a fixed labor cost that increases labor productivity f_l (Murphy, Shleifer, and Vishny (1988a)) or as a continuum of such opportunities that makes for a smoothly increasing productivity ($f_{ll} >$ 0) (Kiyotaki (1988)). The first approach has the appeal of making very clear the rationale behind increasing returns, but leads to the awkward implication of a discrete jump in aggregate productivity in any "representative agent" model. The second approach avoids this problem and thus probably provides a better approximation for a "representative" production function in the absence of more sophisticated aggregation methods. Although both approaches can lead to multiple equilibria, the first one artificially reduces the social increasing returns region to the single point at which aggregate productivity jumps, and overlooks the equilibrium in this region because it occurs at a discontinuity, as illustrated by equilibrium E^*_2 in figure 2. The second approach is clearly more suitable for the next part of this essay, where we analyze dynamics in the increasing returns region.

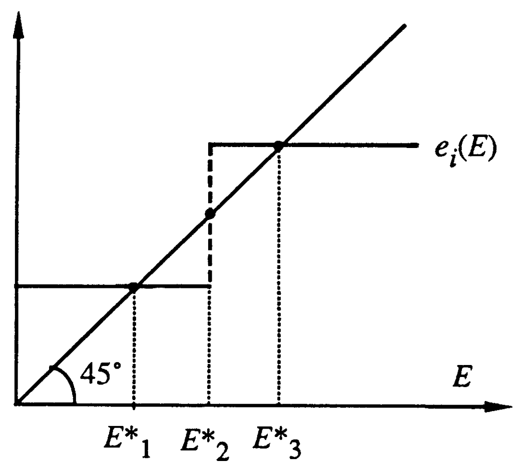


FIGURE 2

Although the ultimate source of social increasing returns in this case is not in itself an externality, its operation relies heavily on the pecuniary externality due to imperfect competition. Indeed, a firm cannot operate in an increasing returns region of its production function unless it considers that the negative effect on price of expanding output outweighs the positive effect on productivity. In other words the SOC must be satisfied and its profit function concave ($V_{ee} < 0$, in terms of the previous section's notation). But if all firms expand output simultaneously, aggregate income Y and their demand functions will rise (all goods are necessarily normal, since they enter U symmetrically), which alleviates the negative effect of expansion on price and allows the positive effect on productivity to increase marginal profits ($V_{ee} + V_{eE} > 0$).

Countercyclical Markups. Finally, multiple equilibria are also possible if the markup $(1+\mu)$ in (1.3.3) falls with aggregate output. In our model this phenomenon arises if the utility function is such that its elasticity of substitution rises with consumption. To see this note that with a large number N of goods, income effects are negligible and, under symmetry, the price elasticity ε of demand for any good i is essentially given by the elasticity of substitution σ between i and any one of the other (symmetric) goods[1]: $\varepsilon \approx \sigma$. If Y units of each good are consumed, this elasticity of substitution will be a function of Y, and can be either increasing or decreasing with Y. If its is increasing, ε will be procyclical and the markup $1+\mu = 1/[1-1/\varepsilon]$ countercyclical.

If the elasticity of demand rises with aggregate consumption, firms will have more incentive to produce when aggregate output is high, hence the possibility of increasing marginal return on effort in the aggregate. It is precisely this mechanism that generates multiple equilibria in Heller (1986), and should not be confused with pure demand spillovers. In principle, this effect does not hinge on the specific reason in this model why markups are countercyclical, and may arise under a countercyclical degree of collusion (Rotemberg and Saloner (1986)) or other countercyclical markup stories discussed in the literature (see Stiglitz (1984)).

The social increasing returns condition in this example shows how little the strategic complementarity inherent in aggregate demand spillovers can achieve in itself in terms of multiple equilibria. It is at best an important auxiliary to the true sources of multiplicity outlined above. Grouping these three sources in the same model makes their complementarity very clear: Strongly decreasing marginal products, or procyclical markups would make it harder for trade externalities to lead to multiplicity, for example. Evidence of a flagrant violation of any one of these ingredients would undermine the potency of the others. This is also true of the requirement for relatively rigid real wages discussed above,

[1]Demand faced by any producer is the result of a tradeoff between essentially two types of goods: his and all the other (symmetric) goods. Therefore, the standard two-good Slutsky equation for the price elasticity of demand $\varepsilon(c)$ for any good i holds:

$$\varepsilon(c) = \alpha\eta(c) + (1-\alpha)\sigma(c),$$

where $\eta(c)$ is the income elasticity of demand for i, $\sigma(c)$ the elasticity of substitution between i and any one of the other (symmetric) goods, and α the share of income spent on good i. For large N, α and the income effect $\alpha\eta(c)$ become negligible. So the elasticity of demand is essentially given by the elasticity of substitution: $\varepsilon(c) \approx \sigma(c)$. With a continuum of firms this equality is exact.

which is an important ingredient in any story although it does not in itself generate multiplicity in this model. But so far the evidence on productivity (Hall (1988), Pratten (1971), Ramey (1987) find evidence of increasing returns, Bils (1987) finds the opposite), markups (Rotemberg and Saloner (1986), e.g.), and real wages (Bils (1985), Geary and Kennan (1985) find little cyclical pattern, but see Kydland and Prescott (1988)) at least does not undermine the kind of effects discussed here.

2. INSTABILITY

I now turn to the dynamic implications of the fact that multiplicity requires a region where returns are increasing in the aggregate. Recall that any two adjacent equilibria differ in that the return on effort V_e is increasing in the aggregate at one of them, and decreasing at the other. The first case corresponds to the representative agent's reaction curve cutting the 45° line form below, the second from above. This is because in, say, the first case with social increasing returns case, each agent would react to an increase in aggregate activity by increasing his activity level *more* that proportionally ($de_i/dE > 1$), for if he were to increase it just proportionally he would end up with a positive marginal return on effort.

This difference in the slope of the reaction curve is what determines the stability properties of different equilibria. Crudely speaking, with a steep reaction curve ($de_i/dE >$ 1), agents tend to "overreact" to increases in aggregate activity and lead to further expansion; with a flat curve ($de_i/dE < 1$), they tend to "underreact" and return to lower activity levels. In other words, under increasing returns, an increase in aggregate activity tends to raise incentives and lead to further increases; under decreasing returns, it tends to lower incentives and get reversed. Thus of two adjacent equilibria, the one with social increasing returns tends to be unstable and the one with decreasing returns tends to be stable.

In what follows I formalize this intuitive argument. The next four sections, in turn, give a dynamic extension of the general model of part 1, define an equilibrium for this economy, formulate the concept of social increasing returns, and analyze its destabilizing effect. Three examples are presented in this part's last section.

2.1. *A Dynamic Economy*

I extend the static economy of section 1.1 to a general dynamic economy. At any instant t, each of the economy's $i = 1,\ldots, N$ agents derives a net payoff V which depends on his "stock" variable s_i and his net "flow" variable $\phi_i \equiv ds_i/dt$, and possibly on their respective averages S and Φ for the economy as a whole. The pair (s, ϕ) may represent such diverse things as the capital stock and net investment, employment and the hiring rate, inventories and inventory investment, the stock of durables and durable purchases, etc. But for convenience, they will often simply be referred to as "capital" and "investment." Agents infinitely lived and discount their payoff at rate $r > 0$. A representative agent must choose the path of $\{\phi(t)\}$ that maximizes:

$$\int_0^{\infty} V\big(s(t), \phi(t), S(t), \Phi(t)\big)e^{-rt}dt$$

$$\text{s.t.} \quad \dot{s} = \phi, \quad V_s > 0, V_\phi < 0, V_{ss}, V_{\phi\phi} \leq 0, V_{\phi s} \leq 0. \tag{2.1.1}$$

V is increasing in s (the marginal "benefit" of "capital" is $b \equiv V_s > 0$) and decreasing in ϕ (the marginal "cost" of "investment" is $c \equiv -V_\phi > 0$). For regularity, I also assume V concave in (s, ϕ) and $V_{\phi s} \leq 0$[1].

This setup allows for two different ways in which "external effects" may enter agents' payoffs: either through the aggregate capital stock S, or through aggregate investment Φ. In the former case, it is the stock variable s that correspond to the "effort" of the static model of section 1.1; in the latter, it is the flow variable ϕ. As will be shown, this corresponds to two conceptually different ways in which strategic complementarities may operate in the economy. In the first case, complementarities require that a greater aggregate level S of "capital" *increase its marginal return b/c*; in the second, they require that a greater aggregate level Φ of investment *decrease its marginal cost c*. In other words, in the first case individual "investment" is more productive in periods when the aggregate "capital" stock is high, in the second "investment" is cheaper in periods of high aggregate "investment."

---

[1]This last condition states that accumulating "capital" is not a way to reduce the marginal cost of net "investment" ($-V_\phi$). It is satisfied in most well-known models, where the only cross-effect of "capital" on the cost of *net* "investment" comes from depreciation and usually works in the right direction.

2.2. *Equilibrium Paths*

In perfect foresight equilibrium each agent i choses his optimal "investment" path $\{\phi_i(t)\}$, taking as given the path of aggregates $\{S(t),\Phi(t)$. A representative agent's optimal investment policy is characterized by the following first order conditions (plus an omitted transversality condition), the SOC being guaranteed by the concavity assumption. At any instant t,

$$\begin{cases} p = c, & c \equiv -V_\phi \\ rp = b + \dot{p}\,, & b \equiv V_s, \end{cases} \tag{2.2.1}$$

where p is the "shadow price" of an existing unit of s. The first equation determines the level of "investment" ϕ, which must be high enough so as to equate the marginal cost $c = (-V_\phi)$ of adding a unit of s to its price. The second equation is an "arbitrage condition" on the path of p, along which the required return rp on holding a marginal unit of s must always equal the "benefit" $b = V_s$ derived from it plus "capital gains" dp/dt.

Throughout, I focus on symmetric equilibria, where $s_i = S$ and $\phi_i = \Phi$, $\forall i$. To solve for the symmetric equilibrium path of S and Φ, rewrite the arbitrage condition (2.2.1) for a representative agent, taking $p = c$ into account:

$$rc = b + \left(c_s\dot{s} + c_S\dot{S} + c_\phi\dot{\phi} + c_\Phi\dot{\Phi}\right). \tag{2.2.2}$$

Under symmetry ($s = S$ and $\phi = \Phi$) and recalling that $\phi = ds/dt$, this equation can be written entirely in terms of S and its time derivatives:

$$\ddot{S} = \frac{1}{c_\phi + c_\Phi}\left[(rc - b) - (c_s + c_S)\dot{S}\right] \equiv f(\dot{S}, S). \tag{2.2.3}$$

Equation (2.2.3) is a second order differential equation in S. At a steady state, aggregate investment must be zero and the marginal return b/c on s must equal the discount rate:

$$\Phi^* = 0 \quad \text{and} \quad r = \left.\frac{b}{c}\right|_{(S^*,\,0,\,S^*,\,0)}. \tag{2.2.4}$$

I will always assume that there exists at least one such steady state.

2.3. *Social Increasing Returns*

In order to analyze the stability properties of a steady state S^* we need to understand the ways in which social increasing returns can arise in this economy. As in the static case, social increasing returns arise when strategic complementarities are strong. But now, as discussed in section 2.1, there are two ways in which strategic complementarities can operate in this economy—through the aggregate "capital" stock or aggregate "investment." Each of these effects, if strong enough, can independently produce a phenomenon of akin to the "social increasing returns" of the static model. They are discussed in turn.

External Effects of "Capital". Consider first the case where strategic complementarities arise through external effects of the aggregate state variable S. By analogy with the static case, define agent i's "reaction curve" $s_i^*(S)$ as the steady-state level that s_i would converge to if the aggregate "capital" stock is set at a constant S from now on. The dynamics of s_i are given by (2.2.2), which can be rewritten taking the constancy of aggregate S ($d\Phi/dt = dS/dt = 0$) into account:

$$\ddot{s}_i = \frac{1}{c_\phi}\left[rc - b - c_s \dot{s}_i\right].$$

One can easily show that the optimal path of $\{s_i\}$ converges along a saddle path to the steady state s_i^* that equates the marginal return on investment to the discount rate:

$$r = \frac{b}{c}\bigg|_{(s_i^*,\, 0,\, S,\, 0)}. \tag{2.3.1}$$

Therefore the slope of the reaction curve can be obtained by totally differentiating (2.3.1):

$$\frac{ds_i}{dS} = -\frac{b_S - rc_S}{b_s - rc_s}. \tag{2.3.2}$$

As in the static case, a symmetric steady state equilibrium S^* for the economy occurs at the intersection of the representative agent's reaction curve with the 45° line. The possibility of multiple steady states depends on whether this reaction curve cuts the 45° line from below or not, i.e. whether $ds_i(S^*)/dS > 1$.

Again, this can be interpreted as a social increasing returns condition. Let $R(S)$ be the marginal return on "capital" when all agents in the economy hold a constant stock $s = S$:

$$R(S) \equiv \frac{b}{c}\bigg|_{(S,\, 0, S,\, 0)}. \tag{2.3.3}$$

Taking (2.3.1) into account, we get:

$$R\,'(S^*) = \frac{1}{c}\left[(b_s + b_S) - r\,(c_s + c_S)\right]_{(S^*,\, 0,\, S^*,\, 0)}. \tag{2.3.4}$$

Under our regularity conditions (2.1.1), comparing (2.3.4) with (2.3.2) shows immediately that social increasing returns $R'(S^*) > 0$ is equivalent to a reaction curve slope $ds_i(S^*)/dS > 1$. With no external effects, regularity (2.1.1) implies social decreasing returns: $R'(S^*) < 0$. Increasing returns necessitates a strong enough positive external effect $b_S - rc_S$, i.e. that the aggregate "capital" stock have a positive effect on its marginal benefit or a negative effect on its marginal cost.

External Effects of "Investment". The aggregate flow variable Φ can also give rise to strategic complementarities. In order to define the appropriate "reaction curve" for this effect, it is important to note a major difference between the effects of state and flow variables. Because the stock variable is fixed in the short run, it only made sense to define the reaction curve $s_i^*(S)$ as a *long run* response of individual to aggregate variables. But since the flow variable is the decision variable in the short run, it makes sense now to define the reaction curve as a *short run* response of individual to aggregate variables, *given the present state of affairs*.

For reasons that will be clear below, define the short run reaction curve as the increase $d\phi_i/dt$ in individual "investment" in response to an expected aggregate increase $d\Phi/dt$, given the present values of $s_i(t)$, $\phi_i(t)$, $S(t)$, and $\Phi(t)$. Furthermore, since I am interested in steady state stability, I assume the economy is presently at steady state: $\phi_i(t) = \Phi(t) = 0$ and $s_i(t) = S(t) = S^*$. The slope of this reaction curve can be deduced from (2.2.2):

$$\frac{d\dot{\phi}_i}{d\dot{\Phi}} = -\frac{c_\Phi}{c_\phi}. \tag{2.3.5}$$

Since this slope only depends on the fixed values of $s_i(t)$, $\phi_i(t)$, $S(t)$, and $\Phi(t)$, it is a constant independant of $d\Phi/dt$ and the reaction curve is a straight line.

As usual, a symmetric equilibrium Φ for the economy occurs at the intersection of the individual agent's reaction curve with the 45° line. But now this curve is a straight line and crosses the 45° line only once, which must occur at the origin since the steady state is an equilibrium. As defined, this reaction curve has no room for multiple equilibria because it focuses on steady state behavior where the only equilibrium level of investment is zero[1]. But whether it crosses the 45° line from below or not will still prove to be a criterion for instability and can now be interpreted as a condition of social decreasing "investment" costs.

Define $C(\Phi)$ as the marginal cost of "investment" when all agents in the economy invest $\phi = \Phi$ and hold the steady state capital stock $s = S^*$:

$$C(\Phi) \equiv c\,|_{(S^*,\Phi,S^*,\Phi)}. \tag{2.3.6}$$

Since:

$$C'(\Phi) = c_\phi + c_\Phi, \tag{2.3.7}$$

[1]Had we focused instead on "balanced" growth paths, the external effects of investment could have generated multiplicity with different levels of Φ as in Azariadis and Drazen (1988), but this is beyond the scope of this paper.

it is clear from (2.3.5) and regularity (2.1.1) that the social decreasing cost condition $C'(0) < 0$ at steady state is equivalent to a reaction curve cutting the 45° line from below. Since, by regularity (2.1.1), costs are increasing in the aggregate ($C'(0) > 0$) in the no-externality case, the social decreasing cost condition requires a strong enough negative external effect c_{Φ} of aggregate "investment" on its marginal cost.

2.4. *Steady-State Stability Properties*

We can now turn to the system's stability properties at a steady state S^*. These are obtained by linearizing the differential equation (2.2.3) in the neighborhood of S^*, and solving for the roots of the characteristic equation:

$$\lambda^2 - f_{\dot{S}}(0, S^*)\lambda - f_S(0, S^*) = 0.$$

This equation has in general two roots λ_1 and λ_2, whose product and sum are given by[1]:

$$\begin{cases} \lambda_1\lambda_2 = -f_S(0, S^*) = \dfrac{(b_s+b_S) - r(c_s+c_S)}{c_{\phi}+c_{\Phi}} \\ \lambda_1+\lambda_2 = f_{\dot{S}}(0, S^*) = r - \dfrac{c_S+b_{\Phi}}{c_{\phi}+c_{\Phi}}. \end{cases} \tag{2.4.1}$$

Using the insights of the previous section, (2.4.1) can be rewritten taking (2.3.4) and (2.3.7) into account:

$$\begin{cases} \lambda_1\lambda_2 = \dfrac{R'(S^*)}{C'(0)/C(0)} \\ \lambda_1+\lambda_2 = r - E \end{cases} \quad \text{where } E \equiv \frac{c_S+b_{\Phi}}{C'(0)}. \tag{2.4.2}$$

Let us first examine the implications of (2.4.2) for the standard no-externality case. We know that in this case returns are decreasing in the aggregate ($R'(S^*) < 0$) and costs increasing ($C'(0) > 0$). First, since $R'(S) < 0$ everywhere, the reaction curve (2.3.2) is downward sloping everywhere and steady state is unique. Moreover $\lambda_1\lambda_2 < 0$, which implies that the two eigenvalues are real and of opposite signs, and therefore that the economy has a unique saddle-path equilibrium path that converges to its unique steady state S^* (the transversality condition must be used to eliminate the possibility of non-convergent equilibrium paths).

The presence of external effects can overturn this saddle-path stability result. Note from (2.4.2) that the sign of $\lambda_1\lambda_2$ depends on $R'(S^*)$ and $C'(0)$. It can be positive either (1) if

[1]The derivation of the expression for $\lambda_1+\lambda_2$ takes into account the fact that $b_{\phi} + c_S = V_{S\phi} - V_{\phi S} = 0$.

returns are increasing with s in the aggregate ($R'(S^*) > 0$), or (2) if "investment" costs are decreasing with ϕ in the aggregate ($C'(0) < 0$). In both cases the eigenvalues' real parts are of the same sign, which is the sign of their sum $\lambda_1+\lambda_2 = r-E$. If the term E is nil (or negative) then $\lambda_1+\lambda_2 > 0$, the eigenvalues' real parts are positive, and S^* is totally unstable. Thus, if the effects measured by E are separate from those that give rise to $R' > 0$ (or $C' < 0$), then in their absence social increasing returns (or decreasing costs) have an essentially *destabilizing* effect on dynamics.

In what sense are the effects measured by E separate? Looking at the nominator of

$$E = (c_S + b_\phi)/C',$$

we find that E measures the "cross" external effects of aggregate S on the marginal cost of ϕ ($c_S = -V_{\phi S}$) and those of aggregate Φ on the marginal benefit of s ($b_\Phi = V_{s\Phi}$). These can be contrasted with the "own" external effects of S on V_s (i.e. b_S) and of Φ on $-V_\Phi$ (i.e. c_Φ), which are mainly responsible for social increasing returns (b_S in $R'(S^*) = [(b_s - rc_s) + (b_S - rc_S)]/c$) and decreasing costs ($c_\Phi$ in $C' = c_\phi + c_\Phi$). The only cross effect that can potentially play a role in generating social increasing returns is $c_S < 0$, i.e. if "investment" costs decrease for a higher aggregate "capital" stock. But I do not know of interesting economic examples of this (the examples of $R' > 0$ in the next section all rely on the own effect b_S), and even if this effect arises it would tend to make E negative and $\lambda_1+\lambda_2 = r - E$ positive, further destabilizing the equilibrium S^*. It is therefore fair to say that the own effects that are responsible for social increasing returns (or decreasing costs) are separate from the cross effects measured by E, and thus that the former have *per se* a destabilizing effect.

But what if there are cross effects that make $E \neq 0$? Assume social increasing returns (or decreasing costs). Then our steady state instability result remains unchanged unless $E > r$, i.e. E is a large enough positive externality. In this case $\lambda_1+\lambda_2 = r - E$ becomes negative, the two eigenvalues' real parts turn negative, and equilibrium becomes stable (this phenomenon was first noted by Howitt and McAfee (1988a)). But this is a different form of stability than the saddle-stability of the no-externality case: S^* is now totally stable and there is a continuum of possible equilibrium paths that converge to it starting from any neighboring S. This generates indeterminacy and leaves room for an independent effect of "animal spirits" (see Woodford (1984)). Thus there is a sense in which the cross external effects that make for a positive E puts "limits" on the economy's expansion and have a stabilizing effect that counteracts the destabilizing effect of increasing returns. In fact, even if these limits on expansion are not strong enough to stabilize S^* (i.e. $E > 0$ but too small), section 3.2 will show that they can still prevent explosive behavior and keep the economy fluctuating near its unstable steady state.

2.5. *Examples*

This section presents three examples of the general model (2.1.1). The first example focuses on firms' hiring decisions, and is one where $R' > 0$. The second focuses on households' durable investment decisions and is one where $C' < 0$. The last example examines capital investment decisions and allows for both possibilities, $R' > 0$ or $C' < 0$. Either one of the three sources of social increasing returns discussed in section 1.3 (external economies, increasing returns in production, or countercyclical markups) can produce the necessary external effects in the following examples. But, for simplicity, I choose to rely on external economies everywhere. I also show in each example how stabilizing limits on expansion can arise that make $E > 0$ and counteract the destabilizing effect of $R' > 0$ or $C' < 0$.

Example 1. The first example is of an economy where external effects enter through the aggregate state variable S, and draws on Howitt and McAfee (1988a). In this example the N agents of our general model (2.1.1) are identified with firms and the state variable $s(t)$ as a representative firm's employment level at t. From this perspective a hiring decision is viewed as a kind of "investment," for which the firm may incur costs related to searching for new workers, training them, supplying them with equipment, and committing to (explicit or implicit) contracts with them. Although somewhat unusual, this view captures an important aspect of the labor market and seems quite promising.

For a representative firm, the costs of increasing the labor force can simply be formalized as an increasing function of the *gross* hiring rate

$$g = g(\phi + \delta s), \quad g', g'' > 0, \tag{2.5.1}$$

where $\phi(t) \equiv ds/dt$ is the *net* hiring rate and δ is the exogenous quit rate. This approach is formally identical to the adjustment cost theory of investment (see Lucas (1967), e.g.), and as usual these costs are assumed convex in view of satisfying the SOC.

Assume all firms are competitive producers of a single good, and that aggregate output or, equivalently, aggregate employment $S(t)$ has a positive effect on their labor productivity a because of "thick-market" economies. Individual output $y(t)$ is given by:

$$y = a(S)s, \quad a' > 0.$$

Output is consumed and labor supplied by M identical households. The representative household consumes $c^H(t)$, works $s^H(t)$ hours, and discounts at rate r the instantaneous utility

$$U = c^H - vs^H.$$

This implies that the interest rate will be r and the real wage rate $w = v$.

It follows that a representative firm's profit maximization problem is:

$$\underset{\{\phi(t)\}}{Max} \int_0^\infty \left[a(S)s - vs - g(\phi + \delta s)\right] e^{-rt} dt \quad \text{s.t.} \quad \dot{s} = \phi. \tag{2.5.2}$$

It is clear that this is a special case of the general model (2.1.1) with

$$V(s, \phi, S, \Phi) = \left[a(S)s - vs - g(\phi + \delta s)\right].$$

Social increasing returns can arise in this model because the fact that labor productivity rises with aggregate output implies that $b_S = V_{sS} = a' > 0$ and can make $R' > 0$.

In the model as developed so far there are no "cross" external effects and $E = 0$. So the economy's steady state, if it has one, is unstable under social increasing returns. But the model can be enriched with limits to expansion. One possibility relies on the finiteness of labor endowments: as the economy expands the pool of unemployed workers shrinks and they become costlier to find (see Howitt and McAfee (1987)). This can be formalized by making hiring costs in (2.5.1) also depend positively on aggregate employment S:

$$g = g(\phi + \delta s, S),\ g_S > 0. \tag{2.5.3}$$

These expansion diseconomies give rise to a cross external effect $c_S > 0$ that makes $E > 0$ and has a stabilizing effect on dynamics.

Example 2. In this second example external effects enter essentially through the aggregate flow variable Φ. I draw on Murphy, Shleifer, and Vishny (1988b) and identify the N agents of the general model (2.1.1) with households, and the state variable $s(t)$ with the representative household's stock of durables.

The single durable good is produced using labor by many identical competitive firms whose labor productivity a depends on average (per capita) aggregate output Y. If $y(t)$ and $l(t)$ are a typical firm's output and employment, then:

$$y = a(Y)l, \quad a' > 0.$$

Note that the real wage rate must be $w = a(Y)$ if there is any employment at all.

A representative household discounts at rate r the instantaneous utility

$$U = u(s) - vl^H, \quad u', v > 0, u'' < 0, \tag{2.5.4}$$

where $l^H(t)$ is individual labor supply. Since all households are identical, there will be zero net borrowing in equilibrium and we can ignore credit markets. Thus a typical household's labor income $wl^H = a(Y)l^H$ is spent entirely on gross investment in durables $\phi + \delta s$,, where $\phi(t) = ds/dt$ is *net* durable investment and δ is the depreciation rate of durables. In other words, in order to invest at the net rate ϕ a household will work

$$l^H = \frac{\phi + \delta s}{a(Y)}. \tag{2.5.5}$$

Since aggregate per capita income Y must also be spent on aggregate gross investment $\Phi+\delta S$, we can substitute the latter for the former in (2.5.5):

$$l^H = \frac{\phi + \delta s}{a(\Phi + \delta S)}.$$

Substituting this expression for labor supply in the utility function (2.5.4), we get the household optimization problem:

$$\underset{\{\phi(t)\}}{Max} \int_0^\infty \left[u(s) - v \frac{\phi + \delta s}{a(\Phi + \delta S)} \right] e^{-rt} dt \quad \text{s.t.} \quad \dot{s} = \phi. \tag{2.5.6}$$

This is clearly a special case of our general problem (2.1.1) with

$$V(s, \phi, S, \Phi) = \left[u(s) - v \frac{\phi + \delta s}{a(\Phi + \delta S)} \right].$$

This example can give rise to social decreasing costs $C'(0) < 0$ because a rise in aggregate durable investment Φ increases productivity and real wages $w = a$, and can thus decrease the utility cost (in terms of foregone leisure) of durable investment. Formally the external effect of aggregate investment on its cost is $c_\Phi = -va'/a^2$, which is negative (because of $a' > 0$) and can therefore make $C' < 0$.

Although there are cross external effects in this economy due to depreciation, one can show that they cancel out and $E = (c_S + b_\Phi)/C' = 0$. One way to modify the model to make $E > 0$ is to make the somewhat unappealing assumption that durables s and leisure $(-l)$ are substitutes in a more general utility function $U = U(s, l)$ than (2.5.4). In this case ($U_{sl} > 0$), $E = -(\delta S^* a'/a^2) U_{sl}/C'(0) > 0$ when $C' < 0$ and can have a stabilizing effect.

Example 3. This last example is concerned with capital investment decisions (s now stands for "capital equipment") and allows for either $R' > 0$ or $C' < 0$, depending on whether external economies are economy-wide, or limited to the capital equipment industry. Unfortunately, the simultaneous presence of these two effects results in some complexity.

As in example 1, there are M households each of whom consumes $c^H(t)$ of the economy's only consumption good at t, supplies $l^H(t)$ units of labor, and discounts $U = c^H - vl^H$ at rate r. The interest rate is therefore r and the real wage rate $w = v$.

The consumption good is produced by N symmetric competitive firms. A representative firm uses $l(t)$ units of labor and $s(t)$ units of capital to produce $y(t)$ units of the good using a CRS production function, which can be written without loss of generality:

$$y = \alpha(S)f(l/s)s, \quad \alpha', f' \geq 0, f'' < 0. \tag{2.5.7}$$

The term $\alpha(S)$ measures a positive external effect of aggregate (consumption good) output (summarized in equilibrium by the aggregate capital stock $S(t)$) on individual firms' productivity.

Equating the marginal product of labor to the wage rate $w = v$:

$$\alpha(S)f'(l/s) = w, \tag{2.5.8}$$

we can solve for firm employment as a function of the individual and aggregate level of capital:

$$l = \lambda(S)s, \quad \lambda_S \geq 0, \tag{2.5.9}$$

where $\lambda(S) \equiv (f')^{-1}[v/\alpha(S)]$. Because of constant real wages, employment is proportional to the capital input s. Because $\alpha' > 0$, it rises in times of high aggregate capital stock S, due to improved labor productivity ($\lambda_S = -v\alpha'/(f''\alpha^2)$ has the sign of $\alpha' \geq 0$).

As far as the capital input is concerned, in order to invest at a gross rate $\phi+\delta s$ ($\phi = ds/dt$) a representative firm must incur the costs $p^K g$, where $p^K(t)$ is the real price of capital equipment and g is the amount of equipment needed. As in example 1, adjustment cost are assumed to be a convex increasing function of gross investment:

$$g = g(\phi + \delta s), \quad g', g'' > 0. \tag{2.5.10}$$

Finally, capital equipment is produced in the capital equipment industry. Each of this industry's many competitive firms can transform $x(t)$ units of consumption good into $k(t)$ units of capital equipment using the simple linear technology:

$$k = \beta(K)x, \quad \beta' \geq 0,$$

where $K(t)$ is the industry's total output (normalized by N for later convenience) and $\beta(K)$ measures an additional external economy specific to this industry. Thus the real price of capital equipment is $p^K = 1/\beta(K) = 1/\beta(g(\Phi + \delta S))$ under perfect competition.

We can now show that each of the N producers of consumption goods faces a value maximization problem that is a special case of the general problem (2.1.1). From (2.5.7) and (2.5.10), a representative firm's instantaneous profits are

$$V = \alpha(S)f(l/s)s - wl - p^K g(\phi + \delta s).$$

Recalling that the real interest rate, wage rate, and price of capital are respectively r, v, and $1/\beta(g(\Phi + \delta S))$, and using the employment determination equation (2.5.9), we get the present value of profits as a special case of (2.1.1):

$$\int_0^\infty V(s, \phi, S, \Phi)e^{-rt}dt = \int_0^\infty \left[\alpha(S)f(\lambda(S))s - v\lambda(S)s - \frac{g(\phi+\delta s)}{\beta(g(\Phi+\delta S))}\right]e^{-rt}dt.$$

While the economy-wide external economies $\alpha' > 0$ operate through the aggregate capital stock S and can generate social decreasing returns ($R' > 0$), the economies $\beta' > 0$ that are specific to the capital equipment industry operate through aggregate investment Φ and can generate social decreasing costs ($C' > 0$). In the former case a greater aggregate capital stock makes capital more productive, in the latter a greater level of aggregate investment makes investment cheaper. To see this consider in turn the case when each of these externalities is operating alone.

If external economies are economy-wide ($\alpha' > 0$ and $\beta' = 0$), a higher aggregate capital stock S increases capital productivity (by (2.5.8),$b_S = \alpha' f > 0$) and can therefore yield social increasing returns ($R' > 0$). In this case there are no cross external effects, $E = 0$ and the economy's steady state (if it exists) is unstable. But, as in example 1, a "congestion" externality that makes investment costs rise with the aggregate capital stock ($g_S > 0$ as in (2.5.3)) because of fixed resources, could make $E > 0$ and have a stabilizing effect on dynamics.

If external economies are specific to the capital equipment industry ($\alpha' = 0$ and $\beta' > 0$), greater levels of aggregate investment Φ increase productivity in that industry, make capital investment cheaper ($c_\Phi = -(g'/g)^2\beta' < 0$), and can therefore make investment costs fall in the aggregate ($C' < 0$). As in example 2, there are cross external effects in this case due to depreciation, but they cancel out and $E = 0$.

3. BUSINESS CYCLES

3.1.*General Discussion*

In response to our instability result, the question that naturally comes to mind is (see Howitt and McAfee (1988a)): In what sense can social increasing returns be of empirical relevance if dynamics tend to be unstable wherever they hold? A general message can be gleaned from the way this question has been dealt with in the literature: Fluctuations may be observed around the unstable region if there are natural *limits to expansion* that form a counteracting force and restrain the stabilizing effects of increasing returns. There are at least three ways of thinking about these limits.

(1) The first builds on the stabilizing effects of cross externalities discussed in section 2.4. It can be shown that even if these "expansion diseconomies" ar too weak to stabilize the economy's steady state, they can still prevent explosive behavior and keep the economy fluctuating near its unstable steady state.

(2) The second approach relies, as a stabilizing factor, on the ultimate disappearance of increasing returns. If the unstable equilibrium is bordered on both sides by two other equilibria, these will necessarily exhibit decreasing returns and be stable. In this case the economy may be disturbed from one to the other, crossing the entire increasing returns region on the way.

(3) Finally, one can rely on "capacity" constraints on both sides of the unstable equilibrium between which the economy can oscillate (as in Murphy, Shleifer, and Vishny (1988b)). Their role can be simply explained in the context of the static model of part 1. Suppose there are constraints $E^{min} \leq e \leq E^{max}$(where E^{min} is possibly zero) on the effort variable e. Then both the high (E^{max}) and the low (E^{min}) effort equilibria are feasible if the reaction curve $e_i(E)$ is steeper than the 45° line, since in this case $e_i(E^{max}) > E^{max}$ and $e_i(E^{min}) < E^{min}$ as illustrated in figure 3. So there can be a cyclical equilibrium fluctuating between these two levels, coordinated by an extraneous "sunspot" variable. This is not possible under decreasing returns, when the reverse inequalities hold.

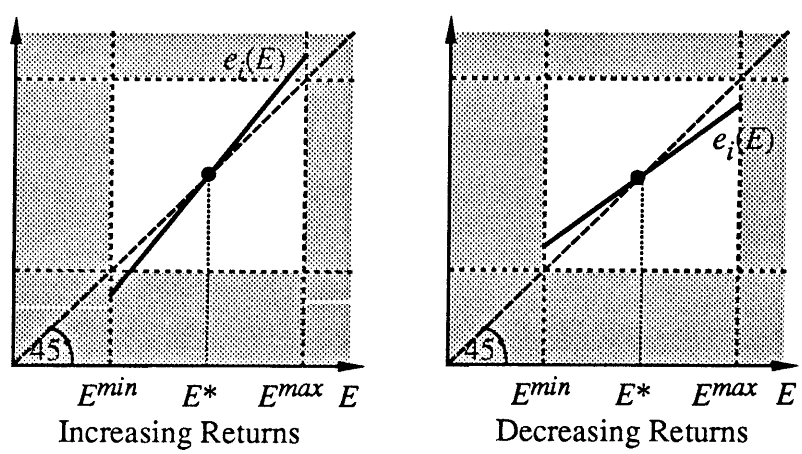


FIGURE 3

There is, however, an unrealistic ring to these last two stories of simultaneous swings from extreme to extreme across the whole increasing returns region. This is partly due to the absence of intertemporal links, and probably partly to the "representative agent" nature of the model, whereas a truly aggregative model where agents have different reaction curves, or different capacity constraints is likely to generate smoother dynamics within the social increasing returns region.

These three ways of thinking about limits to expansion are discussed in the next three sections. There are probably other possible ways. But if the economy is truly fluctuating in an increasing returns region, what the binding constraints on expansion are that maintain it there becomes an empirical question. The essential point here is that such behavior is possible. In this case, multiple equilibria are to be taken seriously, an can potentially provide a partial explanation for observed macro instability.

3.2. *Expansion Diseconomies*

Assume that the economy has a steady state S^* that exhibits either increasing returns $R' > 0$ or decreasing costs $C' < 0$. Recall from section 2.4 that S^* will be totally stable or unstable depending on whether the cross external effect E is greater or smaller than the discount rate r. A positive E was interpreted as measuring diseconomies of expansion which have a stabilizing effect on dynamics. The exact nature of the mechanism behind these limits to expansion depends on the specific model in question. In the examples of section 2.5, a positive E was due to such diverse factors as increasing hiring costs in periods of high employment, the substitutability of leisure and durables consumption, or congestion in capital investment due to fixed resources. To cite another example, in Diamond and Fudenberg (1987) economic activity is measured by the level of inventories and expansion diseconomies arise because inventories tend to get depleted faster as aggregate activity rises. More generally, a positive E can be usefully taken as a proxy for limits to expansion that are likely to arise in richer, more realistic models with heterogeneous agents, money, etc.

It will be convenient to define a parameter ξ that measures the strength of expansion diseconomies: the greater ξ is, the larger E is ($\partial E/\partial \xi > 0$). The critical value of ξ is the value ξ^* at which $E = r$. If $\xi < \xi^*$ equilibrium is unstable, and if $\xi > \xi^*$ it is stable. In the latter case, S^* is totally unstable, and there is a continuum of paths that converge to it.

Steady state stability does not only require some form of expansion diseconomies, but also that their effect be so strong as to overcome the destabilizing effect of increasing returns. Although one may find the presence of some limits to expansion plausible, it is not clear whether these are powerful enough to stabilize S^*. But even if expansion diseconomies are not so powerful ($\xi < \xi^*$) and S^* is unstable, there is still a sense in which their mere presence has a stabilizing effect on the economy. One can show that the presence of some limits to expansion can prevent explosive behavior and keep the economy fluctuating in a "limit cycle" around its otherwise unstable equilibrium (see Diamond and Fudenberg (1987), Drazen (1988), Hammour (1988)).

The argument relies on the Hopf bifurcation theorem (stated formally in the appendix). The theorem concerns dynamic systems whose characteristics may depend on a "bifurcation parameter" ξ. If the linearized system near a steady state S^* has a conjugate pair of imaginary eigenvalues whose real part changes from positive to negative as ξ crosses the critical value ξ^*, then a "bifurcation" occurs at ξ^* and a "limit cycle" will appear around S^* for a neighborhood of ξ on one side of the critical value.

In order for this theorem to apply in our case, we need to check that the linearized system's eigenvalues at S^* are imaginary, and that their real part changes sign at the critical values $\xi = \xi^*$. To see this, note from (2.4.2) that when $\xi = \xi^*$ the sum $\lambda_1+\lambda_2 = r - E = 0$. Since under increasing returns $\lambda_1\lambda_2 > 0$, this means that we have a purely imaginary pair of eigenvalues. By continuity, the eigenvalues will also be imaginary in a neighborhood of ξ^*, with positive real part when $\xi < \xi^*$ and negative when $\xi > \xi^*$. Thus a bifurcation occurs at $\xi = \xi^*$.

Moreover, under a certain stability condition (given in the appendix), the limit cycle in question will be stable, as illustrated in figure 4, and appears for $\xi < \xi^*$, when expansion diseconomies are not powerful enough to stabilize the equilibrium. Thus, even if they are weak, the mere presence of these diseconomies can prevent explosive behavior and maintain the economy fluctuating in the increasing returns region around its unstable equilibrium. In the presence of a stable limit cycle, there is a continuum of equilibrium paths that converge to the cycle. This indeterminacy allows "animal spirits" to play an independent role.

Finally, note that even if there are no expansion diseconomies and $E = 0$, the discount rate r could serve as a bifurcation parameter. Indeed, from (2.4.2) we know that in this case $\lambda_1+\lambda_2 = r$, and the two eigenvalues cross the imaginary axis as r crosses the critical value $r^* = 0$ and goes from negative to positive (although only positive values of r are economically meaningful). Although this approach seems attractive because of its generality, it can be problematic in many economic examples. This is because the equilibrium $S^* = S^*(r)$ usually depends on the bifurcation parameter r, so the Hopf theorem requires that it be well defined on both sides of the critical value $r^* = 0$. Unfortunately, in many specific economic models S^* is only well defined for positive values of the discount rate r.

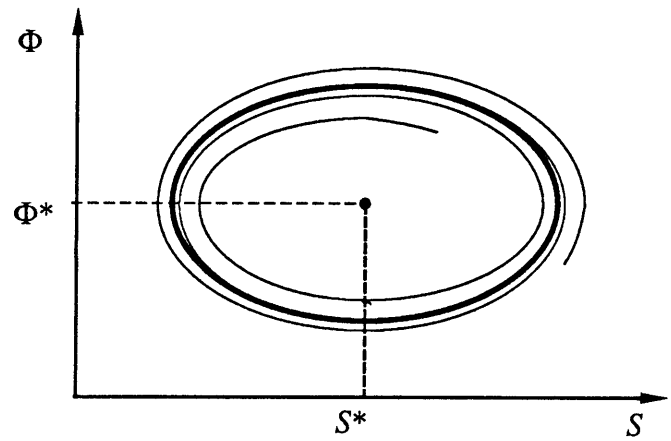


FIGURE 4

3.3. *Multiple Steady States*

Another way of thinking about limits to expansion relies on the ultimate disappearance of increasing returns away from the unstable steady state, and the appearance of adjacent steady states with locally decreasing returns. In this section, I limit myself to increasing returns $R' > 0$ as a source of instability, and assume at any steady state S^* that "investment" costs are well behaved with $C'(0) > 0$, and that there are no strong congestion externalities ($E < r$).

The idea can be developed using a phase diagram representation of the dynamics in (S, Φ)-space. From (2.2.3), these dynamics are governed by:

$$\begin{cases} \dot{S} = \Phi \\ \dot{\Phi} = f(\Phi, S). \end{cases} \tag{3.3.1}$$

As illustrated in figure 5, the ($dS/dt = 0$) schedule is flat, given by $\Phi = 0$. Multiple steady states occur if it is crossed by the ($d\Phi/dt = 0$) schedule more than once. For any steady state S^*, the slope of the ($d\Phi/dt = 0$) schedule is obtained by totally differentiating $f(\Phi, S) = 0$, and can be expressed in terms of the sum and product of the linearized system's eigenvalues at S^* given by (2.4.1):

$$\left.\frac{d\Phi}{dS}\right|_{(\dot{\Phi}=0)} = -\frac{f_S}{f_\Phi} = \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}.$$

Using (2.4.2), we get the following economic interpretation for this slope:

$$\left.\frac{d\Phi}{dS}\right|_{(\dot{\Phi}=0)} = \left(\frac{1}{r-E}\right)\left(\frac{R'(S^*)}{C'(0)/C(0)}\right). \tag{3.3.2}$$

Under our assumptions ($C'(0) > 0$, $r - E > 0$), the ($d\Phi/dt = 0$) schedule is upward sloping if an only if $R'(S^*) > 0$, i.e. under increasing returns. Thus the middle equilibrium S^*_2 in figure 5 exhibits increasing returns and is unstable, and the two lateral equilibria S^*_1 and S^*_3 exhibit decreasing returns and are saddle-path stable. The two saddle paths that converge to the stable equilibria are shown in the figure. (Although I choose to discuss a three steady state example, similar issues can arise with only two steady states).

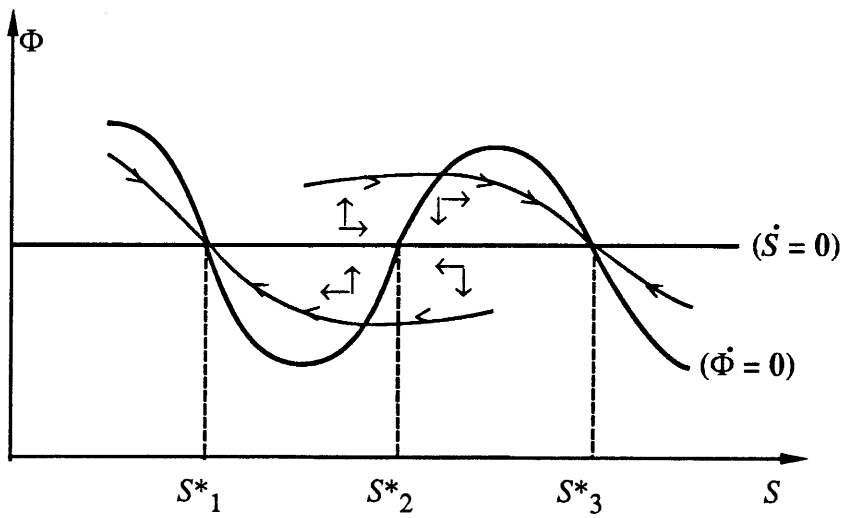


FIGURE 5

It is clear that there is always some region for the aggregate "capital" stock around the unstable equilibrium S^*_2 in which the two saddle paths overlap, i.e. in which one cannot determine whether the economy will converge to the low S^*_1 or high S^*_3 equilibrium. Purely expectational "animal spirits"-like phenomena will determine whether it will enter a boom or a recession.

The region where the two saddle-paths overlap could cover the whole interval between S^*_1 and S^*_3, or could be limited to a smaller subinterval (see Krugman (1989)). In the latter case, the economy's equilibrium is determinate once it leaves the region of overlap. The two saddle-paths that lead to S^*_1 and S^*_3 are two possible equilibrium paths: if the economy starts in their region of overlap, or is disturbed into it, it may jump onto either one and converge to the corresponding steady-state.

If overlap is complete, indeterminacy extends to the whole region between the two stable steady-states and dynamics are much richer (similar behavior can occur if overlap is only complete on one side of S^*_2). Note that any path that starts in the region enclosed by the two saddle-paths between S^*_1 and S^*_3 remains there (i.e. this area is "invariant" with respect to the dynamic system), and is therefore an equilibrium. This implies a continuum of equilibrium paths, and, again, a possible role for "animal spirits." Moreover the Poincarré-Bendixon theorem (see, e.g., Hirsch and Smale, pp. 248-251) shows that all paths within the invariant region must converge to one or more *limit cycles* (or to the equilibrium S^*_2 itself, if it were stable). Thus, dynamics in this case are very similar to the those obtained under "expansion diseconomies." But it is an open question whether the reversal in returns on "capital" that produces multiple steady states is sufficient for complete overlap to occur, or whether some degree of congestion $E > 0$ is still necessary.

3.4. *Capacity Constraints*

A third way to think about limits to expansion is in terms of "capacity" constraints, an idea reminiscent of Hicks' (1950) trade cycle model. Such constraints can take various forms, but in order to illustrate the way they may operate, I choose to discuss the simple case of an individual upper-bound S^{max} on the level of the stock variable: $s \leq S^{max}$. Such a constraint can keep the economy fluctuating within the increasing returns region independently of the presence of the expansion diseconomies and multiple steady states discussed in the previous two sections. To show this most clearly, I assume throughout this section that there are no cross externalities ($b_\Phi = c_S = 0$) and focus on a single unstable steady state S^*. Furthermore, I assume that instability is due social increasing returns ($R'(S^*) > 0$) rather than decreasing costs. This is done only to fix ideas, and it is possible to conduct a parallel discussion with similar results in the case of decreasing costs.

With capacity constraints, it will be necessary to analyze the economy's dynamics in (S, p)-space, rather than in the (S, Φ)-space of the previous section. From (2.1.1) and (2.2.1), dynamics are governed under symmetry ($s = S$, $\phi = \Phi$) by:

$$\begin{cases} \dot{S} = \Phi \\ \dot{p} = rp - b. \end{cases} \tag{3.4.1}$$

In order to close this system, we need to express Φ as a function of (S, p). As long as S is below full capacity, this function is given implicitly by the FOC (2.2.1) which equates the marginal cost of "capital" to its price: $p = c$. When full capacity is reached ($S = S^{max}$) net investment Φ cannot be positive and this condition becomes just an inequality $p \geq c$, with equality if the constraint is not binding ($\Phi < 0$). In other words:

$$\begin{cases} p = c, & \text{if } S < S^{max}; \\ \Phi \leq 0 \text{ and } p \geq c \ \text{(with "=" if } \Phi < 0), & \text{if } S = S^{max}. \end{cases} \tag{3.4.2}$$

Let us first characterize the dynamics in the region $S < S^{max}$. Assuming there is a steady state (S^*, p^*) in this region, the difference between the slopes of the $(d\Phi/dt = 0)$ and the $(dS/dt = 0)$ schedules at (S^*, p^*) can be derived from (3.4.1)-(3.4.2):

$$\left.\frac{dp}{dS}\right|_{(\dot{p}=0)} - \left.\frac{dp}{dS}\right|_{(\dot{S}=0)} = \frac{-R'(S^*)C'(0)/C(0)}{(b_\phi + b_\Phi) - r\,(c_\phi + c_\Phi)}.$$

Because of regularity (2.1.1), increasing costs $(C'(0) > 0)$, and no cross externalities $(b_\Phi = 0)$, the denominator in this expression is negative. Thus, the whole expression is positive if and only if returns are increasing in the aggregate $(R'(S^*) > 0)$. In this case the slope of the $(d\Phi/dt = 0)$ schedule is greater than that of $(dS/dt = 0)$ and dynamics are as illustrated in figure 6. Throughout the remainder of this section, I assume that S^{max} is close enough to S^* so as to allows us to focus on a linearization of dynamics around the steady state (S^*, p^*).

The region where the constraint is staisfied is to the left of the vertical line $(S = S^{max})$. Consider the intersection of this line with the $(dS/dt = 0)$ schedule. Below this point the constraint is not binding (equating $p = c$ requires negative net investment Φ) and dynamics are just a continuation of the no constraint region. Above this point the constraint is binding, and dynamics are given by $\Phi = 0$ and $dp/dt = rp - b^{max}$, where

$$b^{max} \equiv b\left(\left(S^{max}, 0, S^{max}, 0\right)\right)$$

This implies that p is decreasing when $p < b^{max}/r$ and increasing otherwise. This is illustrated in the figure. It can be shown that $p = b^{max}/r$ must be at or above the point where the vertical line $(S = S^{max})$ intersects the $(dp/dt = 0)$ schedule, given that $b_\phi + b_\Phi \geq 0$ (because of regularity (2.1.1) and no cross externalities $(b_\Phi = 0)$).

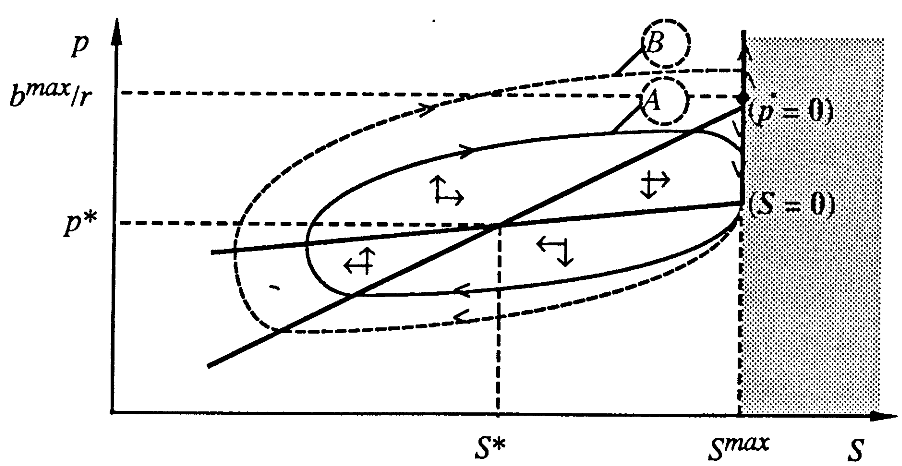


FIGURE 6

The above characterization of dynamics allows us to show the existence of a limit cycle around (S^*, p^*) under certain conditions. Consider the path that goes through the intersection of the line $(S = S^{max})$ and the schedule $(dS/dt = 0)$. Approximating the dynamics by their linearization around (S^*, p^*), we know by steady-state instability that this path must be spiralling outwards. As illustrated in the figure, I distinguish between the case when this path spirals around the equilibrium and meets the line $(S = S^{max})$ below $p = b^{max}/r$ (case A), and when it does not (case B). In the former case, it is easy to see by inspection that this path, closed with a vertical segment of the line $(S = S^{max})$, constitutes a limit cycle to which this economy's continuum of equilibrium paths converge.

But when are we guaranteed that the interesting case A will hold? One can show that it will hold if the discount rate r is small enough. This is because we know from (2.4.2) that the system's eigenvalues are purely imaginary when $r = 0$. This means that the paths of the linearized system are concentric circles and case A must necessarily hold. By continuity, it also holds for a small enough positive r.

4. CONCLUSION

This essay has analyzed macroeconomic phenomena, including multiple equilibria and fluctuations around unstable steady states, that are characteristic of a class of general equilibrium models with "aggregate increasing returns." The presence of such increasing returns requires an "external effect" of aggregates on individual payoffs, otherwise the second order condition for individual optimization would not be satisfied. But the required externality is of a very general nature, and aggregate increasing returns can arise in such diverse situations as an economy where agents search for trading partners in the absence of a Walrasian auctioneer, or an imperfectly competitive economy with increasing returns in production. The individual decision variable that is subject to increasing returns can also be given such diverse interpretations as the level of capital investment, of inventory accumulation, or of hiring. Given this degree of generality, it is important to search empirically for observed economic variables, if there are any, that seem to be subject to aggregate increasing returns. Finding such variables promises to provide a key to explaining the instability of macroeconomic aggregates.

APPENDIX

This appendix gives a formal statement of the Hopf bifurcation theorem based on Guckenheimer and Holmes (1983), pp. 151-2:

Theorem (Hopf): *Suppose that the system* $dz/dt = h_\xi(z)$, $z \in \Re^n$, $\xi \in \Re$, *has an equilibrium* (z^*, ξ^*) *at which the following properties are satisfied:*

(i) $D_z f_{\xi^*}(z^*)$ *has a simple pair of pure imaginary eigenvalues and no other eigenvalue with zero real parts. This implies that there is a smooth curve of equilibria* $(z(\xi), \xi)$ *with* $z(\xi^*) = z^*$. *The eigenvalues* $\lambda(\xi)$, $\overline{\lambda}(\xi)$ of $D_z h_{\xi^*}(z(\xi))$ *which are imaginary at* $\xi = \xi^*$ *vary smoothly with* ξ. *If moreover:*

(ii) $d\mathrm{Re}\lambda(\xi)/d\xi\,|_{\xi=\xi^*} > 0$,

then there exists a periodic solution bifurcating from z^* *at* $\xi = \xi^*$ *and the period of the solutions is close to* $2\pi / |\lambda(\xi^*)|$.

The limit cycle will exist in a neighborhood on one side of the critical value ξ^*. To find out on which side it will appear, the following result can be used (see Guckenheimer and Holmes (1983), p. 152-3; or Marsden and McCracken (1976), pp. 131-5):

Proposition: *Suppose the system* $dz/dt = h_\xi(z)$ *described in Hopf's bifurcation theorem is two-dimensional and that, possibly after a change of coordinates, it can be written in the form:*

$$\dot{z} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = h(z) \quad \text{with} \quad D_z h(z^*) = \begin{bmatrix} 0 & |\lambda^*| \\ -|\lambda^*| & 0 \end{bmatrix},$$

where $|\lambda^*| = |\lambda(\psi^*)|$. *Also let the stability coefficient "a" be defined as:*

$$a = \left[f_{xxx} + f_{xyy} + g_{xxy} + g_{yyy}\right] - \frac{1}{|\lambda^*|}\left[f_{xy}(f_{xx} + f_{yy}) - g_{xy}(g_{xx} + g_{yy}) - f_{xx}g_{xx} + f_{yy}g_{yy}\right].$$

Then if $a < 0$, *the periodic solutions appear when* $\xi > \xi^*$ *and are stable (attracting) limit cycles. If* $a > 0$, *they appear when* $\xi < \xi^*$ *and are unstable (repelling).*

REFERENCES

AZARIADIS, C., AND A. DRAZEN (1988): "Threshold Externalities in Economic Development," mimeo.

BILS, M. (1985): "Real Wages over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy*, 93, 666-689.

——— (1987): "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77, 838-855.

BLANCHARD, O. J., AND L. H. SUMMERS (1988): "Beyond the Natural Rate Hypothesis," *American Economic Review*, 78, 182-187

COOPER, R., AND A. JOHN (1988): "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics*, 103, 441-64.

DIAMOND, P. (1982): "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 90, 881-894.

DIAMOND, P. AND D. FUDENBERG (1987): "Rational Expectations Business Cycles in Search Equilibrium," *Journal of Political Economy*, forthcoming.

DRAZEN, A. (1988): "Self-Fulfilling Optimism in a Trade Friction Model of the Business Cycle," *American Economic Review*, 78, 369-372.

GEARY, P., AND J. KENNAN (1982): "The Employment-Real Wage Relationship: An International Study," *Journal of Political Economy*, 90, 850-871.

GUCKENHEIMER, J., AND P. HOLMES (1983): *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York-Berlin-Heidelberg: Springer-Verlag.

HALL, R. E. (1988): "Increasing returns: Theory and Measurement with Industry Data," mimeo, Stanford University.

HAMMOUR, M. (1988): "Increasing Returns and Endogenous Business Cycles," essay 1 in this thesis.

HELLER, W. P. (1986): "Coordination Failure under Complete Markets with Application to Effective Demand," in *Equilibrium Analysis: Essays in Honor of Kenneth J. Arrow, Vol. 2,* ed. by W. Heller, R. Starr, and D. Starrett. Cambridge: Cambridge University Press.

HICKS (1974): *A Contribution to the Theory of the Trade Cycle*. Oxford: Clarendon Press.

HIRSCH, M. W., AND S. SMALE (1974): *Differential Equations, Dynamical Systems and Linear Algebra*. New York: Academic Press.

HOWITT, P. (1985): "Transaction Costs in the Theory of Unemployment," *American Economic Review*, 75, 88-100.

HOWITT, P. AND R. P. McAFEE (1987): "Costly Search and Recruiting," *International Economic Review*, 33, 89-107.

——— (1988a): "Stability of Equilibria with Trade-Externalities," *Quarterly Journal of Economics,* 103, 261-277.

——— (1988b): "Animal Spirits," mimeo, University of Western Ontario.

KIYOTAKI, N. (1985): "Macroeconomic Implications of Monopolistic Competition," Harvard Ph. D. thesis, Chapter 1.

——— (1988): "Multiple Expectational Equilibria under Monopolistic Competition," *Quarterly Journal of Economics,* 103, 695-741.

KRUGMAN, P. (1989): "History and Expectations," mimeo.

KYDLAND, F., AND E. PRESCOTT (1988): "Cyclical Movements of the Labor Input and Its Real Wage," mimeo, Federal Reserve Bank of Minneapolis.

LUCAS, R. E., JR. (1967): "Adjustment Costs and the Theory of Supply," *Journal of Political Economy,* 75, 321-334.

MARSDEN, J. E., AND M. McCRACKEN (1976): *The Hopf Bifurcation and Its Applications*. New York-Berlin-Heidelberg: Springer-Verlag.

MURPHY, K. M., A. SHLEIFER, AND R. VISHNY (1988a): "The Big Push," *Journal of Political Economy,* forthcoming.

——— (1988b)"Increasing Returns, Durables, and Economic Fluctuations," mimeo, University of Chicago.

——— (1989): "Building Blocks of Market Clearing Business Cycle Models," *NBER Macroeconomics Annual,* forthcoming.

PRATTEN, C. F. (1971): *Economies of Scale in Manufacturing Industry*. London-New York: Cambridge University Press.

PRESCOTT, E. (1986): "Theory Ahead of Business Cycle Measurement," Federal Reserve Bank of Minneapolis, *Quarterly Review*, Fall, 9-22.

RAMEY, V. (1987): "Non-Convex Costs and the Behavior of Inventories," mimeo, San Diego.

ROBERTS, J. (1987): "An Equilibrium Model with Involuntary Unemployment at Flexible, Competitive Prices and Wages," *American Economic Review*, 77, 856-874.

ROTEMBERG, J., AND G. SALONER (1986): "A Super Game Theoretic Model of Price Wars during Booms," *American Economic Review*, 77, 917-926.

STIGLITZ, J. (1984): "Price Rigidities and Market Structure," *American Economic Review*, 74, 350-55.

WEIL, P. (1988): "Increasing Returns and Animal Spirits," mimeo.

WOODFORD, M. (1984): "Indeterminacy of Equilibrium in the Overlapping Generations Model: A Survey," mimeo, Columbia University.

ESSAY 3

ARE BUISNESS CYCLES EXOGENOUS?

The exogenous approach to business cycles traces back the sources of economic fluctuations to the presence external shocks: monetary, fiscal, technological, expectational, or other. The basic underlying assumption under this now paradigmatic approach is that the economy has a stable equilibrium to which it tends to return, so that the only way it can exhibit the observed fluctuations is for it to be hit by exogenous disturbances that keep it away from equilibrium. This stability assumption was clearly stated by Ragnar Frisch (1933) who, together with Slutzky (1937), pioneered this approach:

> ... When an economic system gives rise to oscillations *these will most frequently be damped.* But in reality the cycles we have occasion to observe are generally not damped. How can the maintenance of the swings be explained? ... One way which I believe is particularly fruitful and promising is to study what would become of the solution of a determinate dynamic system if it were exposed to a stream of erratic shocks that constantly upsets the continuous evolution and by so doing introduces into the system the energy necessary to maintain the swings. [Italics added].

The alternative endogenous approach holds that the cycle is the result of a deep structural instability within the economic system. It does not deny that exogenous shocks may have an exacerbating effect, but it holds that fluctuations would persist even in their absence. The underlying assumption here is that the economy's equilibrium—if it has one at all—is unstable, so that the economy does not tend to return to it. The nonlinear multiplier-accelerator models of the '40s and 50's (Hicks (1950), Goodwin (1951)) are a notable example of this approach. More recently, general equilibrium models have been proposed, among others, where increasing returns destabilize the economy's equilibrium and make for endogenous fluctuations (Diamond and Fudenberg (1987), Hammour (1989a, b), Howitt and McAfee (1988), Murphy, Schleifer, and Vishny (1988)).

Because the fundamental difference between these two approaches resides in the stability properties of the economy's internal dynamics, one should be able to discriminate between them on empirical grounds by examining these dynamics. If we see in the data that the economy has tended to converge nearer to its "steady-state" growth path when it was close to it, we should conclude that its equilibrium is stable and that fluctuations are exogenously generated. If, on the contrary, we see that it has tended to diverge away from steady state, we should conclude that the economy's equilibrium is unstable and that cycles are essentially endogenous. This essay proposes to implement this test.

At first sight, one might object to this project, believing that we already know its outcome. Indeed empirical macroeconomists routinely estimate systems governing the cyclical dynamics of economic aggregates and rarely, if ever, encounter the case of steady-state instability. But the problem is that aggregate dynamics are usually estimated within the framework of *linear* time series models, and the assumption of linearity almost entirely precludes the presence in *stationary* series of steady-state instability and autonomous fluctuations. This is so because in a linear dynamic system the equilibrium's local stability properties are the same as the system's global properties: Either the equilibrium is unstable and the dynamics are explosive, or it is stable and the dynamics are stationary. Since stationarity is a defining feature of the "cyclical" component of economic aggregates[1], any linear model of their dynamics will necessarily have a stable steady state. Thus linear time series macroeconometrics as it is commonly practiced implicitly adopts the Slutzky-Frisch exogenous approach.

Once we allow for nonlinearities—evidence of which is documented, e.g., in Brock and Sayers (1988) for several aggregates—these strict limitations can be avoided. Dynamics can be both globally stationary, and locally unstable at steady state. Although exogenous shocks may still play an important role, fluctuations can persist indefinitely in stationary systems even in the absence of such shocks. The simplest case in which this happens is when the dynamics converge to a stable "limit cycle" around an unstable equilibrium. More generally, very complex, even "chaotic" fluctuations may occur whose erratic behavior may owe little to the presence of "random" shocks[2].

In order to discriminate between exogenous and endogenous dynamics its is therefore imperative to use a nonlinear model. One simple class of nonlinear models to work with are stepwise-linear models. Despite their simplicity, they allow for steady-state instability in stationary series, and can exhibit very rich dynamics. In this essay I estimate a stepwise-linear time series model for the growth rate in postwar U.S. industrial production and test for the stability of the system's steady state. Surprisingly, I find that the data rejects the hypothesis of stability. A similar conclusion is reached using the postwar U.S. non-agricultural civilian employment series. If confirmed, this result would suggest that the economy must be undergoing endogenous fluctuations around trend.

These results do not rule out the effects of exogenous shocks. But they do suggest that, beyond these shocks, much of the business cycle may be due to instability in the economy's internal dynamics. This conclusion has important implications on the ultimate sources of instability in the economy. From a theoretical point of view, it may require a paradigm shift from exogenous to endogenous theories of the cycle. From an empirical point of view, it undermines the methodological validity of the standard linear time series macroeconometrics that naturally complement exogenous theories and have dominated the literature in the past fifteen years.

---

[1]See, e.g., Stock and Watson (1988).

[2]For an introduction to nonlinear dynamics from the point of view of bifurcation theory, see Guckenheimer and Holmes (1983).

The rest of this essay is divided into four parts. The first presents the nonlinear time series model that will be estimated and a method for testing for the stability of its steady state. The second two parts estimate the model and implement the test for the industrial production and the employment series, respectively. The last part concludes.

1. A NONLINEAR MODEL

1.1. *A SETAR Model*

Suppose the dynamics of a stationary aggregate series $\{x_t\}$ (the next two parts consider the growth rate in industrial production and employment) are governed by the system:

$$x_t = F(x_{t-1}, x_{t-2}, \dots) + \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is a sequence of exogenous disturbances[1]. Further assume that the deterministic system corresponding to (1) has a unique steady state, defined as a point x^* that satisfies:

$$x^* = F(x^*, x^*, \dots).$$

The question about the exogeneity of the observed fluctuations in x_t can be stated as follows: Suppose the exogenous shocks ε_t are suddenly interrupted, would system (1) still exhibit sustained fluctuations in the long run[2]? If the answer is no, then the observed fluctuations in x_t are due to the exogenous shocks ε_t. As argued in the introduction, this requires that the steady state x^* be stable. If the answer is yes, then fluctuations are essentially due to the internal dynamics of system (1).

Thus one way to test for the exogeneity of fluctuations is to test for the stability of x^*. As explained in the introduction, it is important to allow for nonlinearities in F under which the stability properties at x^* may be different from those away from x^*. Generally speaking, in order to test for steady-state stability, we must estimate a linearization of F in a neighborhood of $(x^*, \dots, x^*)$ and examine its stability properties.The simplest way this can be done is to model F as a piecewise linear function:

---

[1]i.e., ε_t, is independent of the history $(x_{t-1}, x_{t-2}, \dots)$.

[2]This thought experiment is presented primarily because of its intuitive appeal. Formally speaking, it always makes sense. But, as Kevin Murphy pointed out to me, it may not make economic sense. Removing the exogenous shocks ε_t would amount to removing uncertainty from the economy, which in turn could modify agents' behavior as embodied by the function F that describes internal dynamics. Although this shows the caveats of using such a thought experiment, the real issues are unaffected as long as it is not the behavioral effect of uncertainty that destabilizes the economy's steady state.

$$x_t = \begin{cases} \phi_{10} + \phi_{11}x_{t-1} + \phi_{12}x_{t-1} + \dots + \phi_{1k}x_{t-k} + \varepsilon_t & \text{if } x_{t-1} \le r_1; \\ \phi_{20} + \phi_{21}x_{t-1} + \phi_{22}x_{t-1} + \dots + \phi_{2k}x_{t-k} + \varepsilon_t & \text{if } r_1 < x_{t-1} \le r_2; \\ \phi_{30} + \phi_{31}x_{t-1} + \phi_{32}x_{t-1} + \dots + \phi_{3k}x_{t-k} + \varepsilon_t & \text{if } r_2 < x_{t-1}, \end{cases} \quad (2)$$

where it is assumed that this system has a unique equilibrium $r_1 < x^* < r_2$. In this case, the second equation gives the local dynamics at x^*, while the first and third give the (possibly very different) dynamics away from x^*. This kind of model is known in the literature as a SETAR(2; k, k, k) (a Self-Exciting Threshold Auto-Regressive model with 2 thresholds r_1 and r_2, and k lags in each equation) and possesses a well-developed statistical theory (see Tong (1983)).

In order to illustrate the kind of dynamics this model can generate, consider the simple one-dimensional case ($k = 1$) where $\phi_{10} = \phi_{20} = \phi_{30} = 0$ and $r_1 < 0 < r_2$, so that $x^* = 0$ is the system's unique steady state. In this case it is quite possible for x^* to be unstable ($|\phi_{21}| > 1$) and for the dynamics around steady state to remain stationary because of "stability" away from it ($|\phi_{11}|$, $|\phi_{31}| < 1$). Figure 1 gives a simulated example of the resulting deterministic internal dynamics (ε_t is set to zero). Despite the system's instability, the estimation of a (misspecified) linear model using 500 simulated data points gives:

$$x_t = 0.04 - 0.25x_{t-1}.$$

with the misleading impression of steady-state stability ($-1 < -0.25 < 1$).

1.2. *Testing for Exogeneity*

The exogeneity test described in the previous section is a test of steady-state stability. The maintained hypothesis is that the economy has a unique steady state x^* in the range of its observed fluctuations, and that $r_1 < x^* < r_2$ (this will be tested for and verified in section 2.3). Therefore x^* is the equilibrium of the second equation of system (2):

$$x_t = \phi_{20} + \phi_{21}x_{t-1} + \phi_{22}x_{t-2} + \dots + \phi_{2k}x_{t-k} + \varepsilon_t.$$

It is well known that x^* is stable if and only if all the roots of the characteristic equation:

$$h(z) \equiv z^k - \phi_{21}z^{k-1} - \phi_{22}z^{k-2} - \dots - \phi_{2k} = 0$$

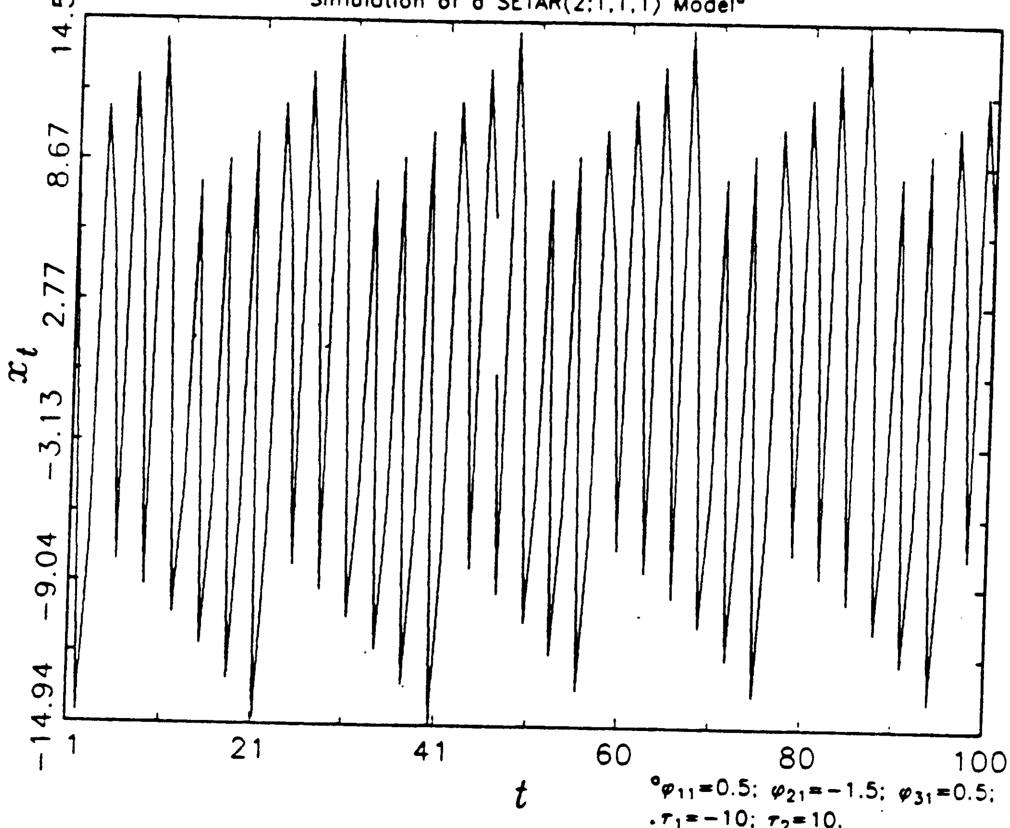
are within the unit circle (see, e.g., Goldberg (1958), pp. 163-4).

Testing for stability by calculating the roots of the characteristic equation would be forbiddingly complex. There are, however, simple implications of the fact that no roots of this equation lie outside the unit circle. First consider the case of positive real roots. The characteristic equation has no real roots greater than one if and only if $h(z)$ does not cross

GAUSS Nov. 29, 1988 10:27:30 PM

FIGURE 1

Simulation of a SETAR(2;1,1,1) Model°



the 0-axis for $z \geq 1$. Since $h(z)$ is continuous and goes to infinity with z, this implies $h(1) > 0$, which can be written as:

$$H_0: \sum_{i=1}^{k} \phi_{2i} < 1. \tag{3}$$

This is the well-known condition that the sum of lag-coefficients should be less than one. If H_0 is rejected, then x^* is unstable and the cycle is endogenous (the converse is not true). By a similar argument, one can show that the corresponding condition for no real roots less than -1 is:

$$\sum_{i=1}^{k} (-1)^i \phi_{2i} < 1.$$

But since it turns out that my data do not reject this condition, I will not come back to it.

2. ESTIMATION AND TESTING: INDUSTRIAL PRODUCTION

This part turns to the implementation of the stability test to the industrial production series. I choose to focus on the U.S. industrial production index in the post-war period (1948:01-1988:04), because of its monthly availability and its high cyclical volatility. I define the series $\{x_t\}$ as the first difference of the logarithm of the industrial production index, and assume it is stationary. x_t can be interpreted as an approximation to the monthly growth rate in industrial production, and averaged 0.31% over the sample.

2.1. *Model Selection*

The SETAR(2; k, k, k) system in (2) describes a class of models with different numbers of lags k and different threshold values r_1 and r_2. One of these models must be selected before estimation can be done. As far as k is concerned, enough lags should be included to insure that most of any eventual moving average structure in the disturbances ε_t is eliminated (see Harvey (1981), p. 36). I choose to regress x on its lagged values over the past year (i.e. $k = 12$), and assume this leaves the ε_t's approximately i.i.d.

Selecting the threshold values is more delicate, and amounts to selecting a functional form for F in (1). The thresholds were introduced in order to isolate the local dynamics around steady state. Thus selection of the interval between r_1 and r_2 should be governed by the following criteria: (1) It should be wide enough and located in such a way as to give us enough confidence that the steady state is between r_1 and r_2; (2) It should be narrow enough so as to allow us to focus on the local dynamics around x^*; (3) It should be wide enough so that enough observations fall between r_1 and r_2 to give us confidence in the parameter estimates. These considerations basically call for subjective judgement. On the other hand, in order to avoid arbitrariness as much as possible, the thresholds should capture a nonlinearity that is present in the data. This can be achieved by more systematic maximum likelihood methods, which must somehow be combined with subjective judgement.

One way of thinking of the subjective criteria is as identifying restrictions that must be satisfied by the threshold values. From this point of view we should select the values of r_1 and r_2 that maximize the likelihood function subject to these restrictions. This can be achieved by finding the constrained maximum likelihood values of r_1 and r_2 assuming:

$$r_1^{\min} < r_1 < r_1^{\max} \quad \text{and} \quad r_2^{\min} < r_2 < r_2^{\max},$$

where the bounds $r_1^{\min}$, $r_1^{\max}$, $r_2^{\min}$, and $r_2^{\max}$ are chosen subjectively so as to satisfy criteria (1)-(3). If an *interior* maximum is found, the selected thresholds capture a natural nonlinearity near steady state and satisfy the identifying restrictions. They need not be equal to the global maximum likelihood values because the latter may be capturing a nonlinearity further away or closer to steady state which satisfies the criteria less well. Using the global values would make the selection process more systematic, but may undermine the meaningfulness of the results.

Assuming the errors are normally distributed, maximizing the likelihood function is almost the same as minimizing the sum of squared residuals (SSR), which can be calculated by estimating the parameters using the method discussed in the next section. Because the SSR function is not "smooth" in r_1 and r_2, I used a grid search to minimize it[1].

The selection procedure was implemented as follows. The average value of x_t over the sample was 0.31%. If the dynamics were linear, this would be the estimate of the system's steady state (see Harvey (1981), p. 123). In the context of our model this is only an approximation. Based on this, I first restricted the threshold values to lie on either side of this approximate value:

$$r_1 < 0.31\% < r_2,$$

since the actual steady state should be between them. Further experimentation shows that maximizing likelihood over the range:

[1] The grid search used increments of 0.01%, and then sometimes finer increments around the minimum value.

$$-0.20\% < r_1 < 0.31\% < r_2 < 0.70\% \tag{4}$$

gives an interior solution $r_1 = 0.19\%$ and $r_2 = 0.65\%$ that is statisfactory in terms of criteria (1)-(3). It is these threshold values that I use for estimation and testing in the next section.

How restrictive are the upper- and lower-bound constraints in (4)? In fact, it is only the upper-bound that turns out to really matter. If r_2 is allowed to be greater than 0.72%, another solution at $r_1 = 0.21\%$ and $r_2 = 0.745\%$ would naturally arise. The corresponding estimation and testing results are presented in the appendix. As shown there, the only problem with these values is that, because of their asymmetry around the average value of x, they seem to make us less confident about the location of the steady state. But the test's conclusions are basically the same, although weaker because of greater stability due to the wider interval between r_1 and r_2.

2.2. *Results*

In what follows I assume that the SETAR(2; 12, 12, 12) model with threshold values $r_1 = 0.19\%$ and $r_2 = 0.65\%$ is a good approximation for our purposes of the "true" model governing the data. Given this, the estimation method is described in chapter 4 of Tong (1983) and consists basically in separating the data points x_t into three groups corresponding to the three equations in (2), depending on the value of x_{t-1}. Each equation is then estimated using standard OLS. The resulting parameter estimates are consistent and asymptotically normal. They are given in table 1. These results should be compared with what would have been obtained using a linear AR model. Using the same number of lags, the estimated AR(12) model for the same data set is:

$$\begin{aligned} x_t \;=\; & 0.002 + 0.388x_{t-1} + 0.075x_{t-2} + 0.046x_{t-3} + 0.031x_{t-4} - 0.052x_{t-5} - 0.021x_{t-6} \\ & + 0.038x_{t-7} + 0.017x_{t-8} - 0.008x_{t-9} - 0.007x_{t-10} + 0.101x_{t-11} - 0.246x_{t-12} + \varepsilon_t. \end{aligned}$$

Table 1 also gives the steady states x_i^* associated with each equation $i = 1, 2, 3$ of model (2), where:

$$x_i^* = \frac{\phi_{i0}}{1 - \sum_{j=1}^{12} \phi_{ij}}. \tag{5}$$

The only estimated equation whose steady state falls within the range for which it is applicable is the second one ($r_1 < x_2^* < r_2$). This is consistent with the maintained hypothesis that system (2) has a unique steady state x^* located between r_1 and r_2, and will be confirmed by more formal tests in the next section. Our estimate of x^* is 0.48%.

TABLE 1

ESTIMATION OF A SETAR(2; 12, 12, 12) MODEL WITH $r_1 = 0.19\%$ AND $r_2 = 0.65\%$: U.S. INDUSTRIAL PRODUCTION, 1947:01-1988:04

	Equation $i = 1$	Equation $i = 2$	Equation $i = 3$
Estimated coefficients:			
ϕ_{i0}	0.002	- 0.006	0.001
ϕ_{i1}	0.332	2.268	0.437
ϕ_{i2}	0.232	0.067	0.031
ϕ_{i3}	- 0.121	0.188	0.106
ϕ_{i4}	0.090	- 0.184	0.058
ϕ_{i5}	0.103	- 0.159	- 0.085
ϕ_{i6}	- 0.145	0.062	- 0.026
ϕ_{i7}	0.060	0.013	0.060
ϕ_{i8}	- 0.019	- 0.048	0.022
ϕ_{i9}	0.021	0.054	- 0.059
ϕ_{i10}	0.054	- 0.018	- 0.050
ϕ_{i11}	0.060	0.166	0.146
ϕ_{i12}	- 0.257	- 0.146	- 0.277
Steady-state of the estimated equation:			
x_i*	0.35%	0.48%	0.19%
Sum of estimated lag-coefficients:			
$\sum_{j=1}^{k} \phi_{ij}$	0.409	2.264	0.363

Finally, table 1 gives the sum of the estimated lag-coefficients for each equation. An exogenous cycle theorist would expect it to be less than one for all three equations. The facts are otherwise. It is smaller than one (0.409 and 0.363) for the first and third equations, and greater than one (2.264) for the second. This is consistent with the idea of local instability near steady state, and global stability away from it. It can be shown that the sum of the second equation's estimated lag-coefficients generally falls as the interval $[r_1, r_2]$ is widened, and tends to the limit 0.360 which corresponds to a linear AR(12) model. This is because the local properties near steady state get blurred as we introduce points further and further away from it. Since the parameter estimates are asymptotically normal, a formal "normal test" of the steady-state stability hypothesis H_0 in (3) can be implemented. The results are presented in table 2.

TABLE 2
TESTS OF STEADY-STATE STABILITY

Number of observations used	Sum of lag-coefficients	Value of test-statistic	P-value
Nonlinear SETAR(2; 12, 12, 12) model[a]:			
115	2.264	2.203	0.014
Linear AR(12) model:			
483	0.360	- 7.613	1.000

[a]Second equation of the model with threshold values: $r_1 = 0.19\%$ and $r_2 = 0.65\%$.

The test in table 2 is based on the 115 observations x_t for which $r_1 < x_{t-1} \leq r_2$. It rejects the stability hypothesis H_0 at the 5% significance level. We can be confident that the steady state x^* is unstable. This should be contrasted with the result that would have been obtained using a linear AR(12) model. The sum of lag-coefficients of the estimated AR(12) model is 0.360 and, to the extent that the stability test is meaningful at all[1], the stability hypothesis can almost never be rejected at any significance level.

2.3. *Checking the Maintained Hypothesis*

The conclusion of steady-state instability and endogenous fluctuations that can be drawn from the above test depends on the maintained hypothesis that the system's unique steady state x^* is between r_1 and r_2. Although we saw that the estimated equations are consistent with this assumption, it would be of interest to test this more formally. Our confidence in this assumption becomes stronger if we can show that equation 2 is the only one to possess a steady state in the range for which it is applicable. In other words, its is stronger if we can *reject* as many of the following as possible: (1) $x_1^* < r_1$; (2.1) $x_2^* < r_1$; (2.2) $r_2 < x_2^*$; (3) $r_2 < x_3^*$.

In order to test any of these hypotheses we need to know whether the sum of lag-coefficients of the appropriate equation is smaller or greater than one. To see why, suppose we want to test whether $x_2^* < r_1$. By (5) this is equivalent to:

[1]Steady-state stability is required for stationarity in linear models, and must be assumed before the test is implemented.

$$\frac{\phi_{20}}{1 - \sum_{j=1}^{12} \phi_{2j}} < r_1. \tag{6}$$

Given the test in table 2, we can confidently assume that the sum of lag-coefficients is greater than one. Under this assumption hypothesis (6) becomes:

$$r_1 < \phi_{20} + r_1 \sum_{j=1}^{12} \phi_{2j},$$

which can be the subject of a normal test if we take r_1 as given by the model. The corresponding tests for the steady states of the first and third equations also require assumptions about the sum of their lag-coefficients. Following the estimates in table 1, I assume both are less than one. The four resulting tests of steady-state location are presented in table 3.

TABLE 3
TESTS OF STEADY-STATE LOCATION[a]

Steady state	Estimate	Hypothesis	Value of test-statistic	P-value
(1) x_1*	0.35%	H_0: $x_1^* < r_1$	0.766	0.222
(2.1) x_2*	0.48%	H_0: $x_2^* < r_1$	2.386	0.009
(2.2) x_2*	0.48%	H_0: $x_2^* > r_2$	1.463	0.072
(3) x_3*	0.19%	H_0: $x_3^* > r_2$	3.179	0.001

[a]Using a SETAR(2; 12, 12, 12) model with threshold values: $r_1 = 0.19\%$ and $r_2 = 0.65\%$.

The only hypothesis in table 3 that cannot be rejected at commonly accepted significance levels is $x_1^* < r_1$. Although this hypothesis cannot be rejected with as much confidence as the others, there is practically no evidence that it holds (it can be rejected at the 23% level). Thus these results generally support the maintained hypothesis behind the stability test, that system (2) possesses a unique steady state x^* located between r_1 and r_2.

2.4. *Simulation*

Note that the estimation of the middle equation of system (2) that gives a local linearization of dynamics around steady state is independent of the specification of the global dynamics away from steady state, as long as the two thresholds r_1 and r_2 are given.

Given these threshold values, we could have given a much richer specification for the dynamics away from equilibrium without affecting the steady-state stability test. But in fact, a very simple piecewise linear specification of the global dynamics was assumed in order to provide a more systematic way of selecting the thresholds.

This very simple specification can be justified by our primary interest in the local characteristics of dynamics around steady state. Any serious attempt to estimate global dynamics that goes beyond the pragmatic concern for a systematic procedure for threshold selection should obviously spend much more effort on specification away from steady state. But it is still interesting to see the kind of endogenous dynamics our very simple global specification can generate.

I simulated the SETAR system estimated in section 2.2 *in the absence of the exogenous shocks* ε_t. The results were disappointing. The simulated time series followed an apparently chaotic path. But, compared with the original industrial production series, the simulated fluctuations had a higher frequency an an extremely low amplitude. Our simple global specification clearly does not capture interesting endogenous dynamics in the industrial production series. A richer specification is needed.

This does not necessarily invalidate the steady-state instability result of section 2.2, which does not hinge on a correct specification of global dynamics, by only on a systematic way of selecting the thresholds between which the dynamics are locally linearized. But it does put into question the meaningfulness of the tests in section 2.3 that are intended to check the maintained hypothesis of a unique steady state, since the latter do depend on an exact specification of global dynamics. Moreover, the question remains open, whether we can find a good nonlinear global specification that captures endogenous dynamics in the industrial production series.

3. RESULTS USING EMPLOYMENT DATA

It is interesting to repeat the same test of part 2 on a series other than industrial production. In this part I examine the U.S. total non-agricultural civilian employment series in the postwar period 1947:01-1988:04. The series is again monthly, and I now define x_t as the first difference of the logarithm of total employment and follow the exact same estimation and testing procedures as in the previous part. It turns out that the results obtained using employment data are similar to those obtained using the industrial production index. This should make us a bit more confident in our result, and in particular makes it much less likely that our conclusion is due to the peculiar way any single series is constructed.

Model selection turned out to be more straightforward than with industrial production. The employment series averaged a 0.17% monthly growth rate over the sample. Maximizing likelihood with the two thresholds on each side of this average value:

$$r_1 < 0.17\% < r_2,$$

gives a solution $r_1 = 0.105\%$ and $r_2 = 0.230\%$ that is satisfactory in terms of the criteria criteria (1)-(3) of section 2.1.

System (2) was estimated using these threshold values, and the results are in table 4. Again, the only estimated equation whose steady state falls within the range for which it is applicable is the second one ($r_1 < x_2^* < r_2$). This is consistent with the maintained hypothesis that system (2) has a unique steady state x^* located between r_1 and r_2, and will be tested shortly. Our estimate of x^* is 0.133%.

The parameter estimates should be compared with the estimated AR(12) model for the same data set:

$$\begin{aligned} x_t \;=\; & 0.001 - 0.036x_{t-1} + 0.172x_{t-2} + 0.116x_{t-3} + 0.076x_{t-4} - 0.024x_{t-5} + 0.052x_{t-6} \\ & + 0.087x_{t-7} - 0.080x_{t-8} + 0.071x_{t-9} + 0.034x_{t-10} - 0.004x_{t-11} - 0.160x_{t-12} + \varepsilon_t. \end{aligned}$$

The sum of the estimated lag-coefficients for this AR(12) model is 0.305, which is less than one. For the SETAR model this sum is smaller than one (0.195 and 0.277) for the first and third equations, and greater than one (2.141) for the second.

TABLE 4

ESTIMATION OF A SETAR(2; 12, 12, 12) MODEL WITH $r_1 = 0.105\%$ AND $r_2 = 0.230\%$: U.S. TOTAL NON-AGRICULTURAL CIVILIAN EMPLOYMENT, 1947:01-1988:04

	Equation $i = 1$	Equation $i = 2$	Equation $i = 3$
Estimated coefficients:			
ϕ_{i0}	0.001	- 0.002	0.001
ϕ_{i1}	- 0.094	2.046	0.046
ϕ_{i2}	0.236	0.065	0.166
ϕ_{i3}	0.147	0.052	0.037
ϕ_{i4}	0.153	0.085	- 0.062
ϕ_{i5}	- 0.051	- 0.256	0.055
ϕ_{i6}	- 0.041	0.090	0.148
ϕ_{i7}	0.069	0.281	0.032
ϕ_{i8}	- 0.004	- 0.085	- 0.155
ϕ_{i9}	0.044	- 0.123	0.151
ϕ_{i10}	0.035	- 0.226	0.102
ϕ_{i11}	- 0.044	0.178	- 0.089
ϕ_{i12}	- 0.255	0.036	- 0.154
Steady-state of the estimated equation:			
x_i^*	0.155%	0.132%	0.148%
Sum of estimated lag-coefficients:			
$\sum_{j=1}^{k} \phi_{ij}$	0.195	2.141	0.277

The formal test whether the sum of lag-coefficients is less than one is presented in table 5. The test based on the 76 observations x_t for which $r_1 < x_{t-1} \leq r_2$ (less than the 115 observations available for industrial production).The stability hypothesis H_0 can only be marginally rejected at the 10% significance level, which is a weaker result than in section 2.2. Only at the 10% level can we be confident that the steady state x^* is unstable. This should, however, be contrasted with the result that would have been obtained using a linear AR(12) model. The sum of lag-coefficients of the estimated AR(12) model is 0.305 and, to the extent that the stability test is meaningful at all, the stability hypothesis can almost never be rejected at any significance level.

TABLE 5
TESTS OF STEADY-STATE STABILITY

Number of observations used	Sum of lag-coefficients	Value of test-statistic	P-value
Nonlinear SETAR(2; 12, 12, 12) model[a]:			
76	2.14	1.255	0.105
Linear AR(12) model:			
483	0.305	- 6.343	1.000

[a]Second equation of the model with threshold values: $r_1 = 0.105\%$ and $r_2 = 0.230\%$.

Finally, table 6 describes the tests concerning the maintained hypothesis that system (2) has a unique steady state x^* that falls between the thresholds r_1 and r_2. Again, as in section 2.3, we can be more confident about the maintained hypothesis the more of the following propositions we can reject: (1) $x_1^* < r_1$; (2.1) $x_2^* < r_1$; (2.2) $r_2 < x_2^*$; (3) $r_2 < x_3^*$. Two of these hypotheses can be rejected at the 5% significance level: $r_2 > x_2^*$ and $r_2 < x_3^*$. The other two hypotheses cannot be rejected at commonly accepted significance levels, but there is no evidence that they hold.

TABLE 6
TESTS OF STEADY-STATE LOCATION[a]

Steady state	Estimate	Hypothesis	Value of test-statistic	P-value
(1) x_1^*	0.155%	H_0: $x_1^* < r_1$	0.910	0.181
(2.1) x_2^*	0.133%	H_0: $x_2^* < r_1$	0.468	0.320
(2.2) x_2^*	0.133%	H_0: $x_2^* > r_2$	1.693	0.045
(3) x_3^*	0.148%	H_0: $x_3^* > r_2$	1.754	0.040

[a]Using a SETAR(2; 12, 12, 12) model with threshold values: $r_1 = 0.105\%$ and $r_2 = 0.230\%$.

In sum, the results obtained using employment data are very similar to those obtained using industrial production. They are however somewhat weaker, since we can only reject the stability hypothesis at the 10% level and we are a little less confident about the maintained hypothesis.

4. CONCLUSION

This essay used a very simple nonlinear time series model to test the exogeneity and steady-state stability of business cycle dynamics. Applying the test to the growth rate in U.S. industrial production, the stability hypothesis was rejected. Intuitively, in periods when the growth rate of industrial production was near its steady state, it was found to diverge away from it rather than converge closer to it. Similar results were obtained using employment data. If confirmed, this result would imply that business cycle dynamics are not due solely to external disturbances, but are essentially endogenous. It would suggest that more research effort should be spent on understanding the unstable nature of the economy's internal dynamics, and perhaps less on the different kinds of external shocks that may be disturbing it. In this respect, the recent models, mentioned in the introduction, of endogenous cycles based on increasing returns are particularly promising. From an empirical point of view, it would put into question the methodological validity of the linear time series macroeconometrics that usually complement exogenous theories.

APPENDIX

This appendix presents estimation and testing results for industrial production using the thresholds $r_1 = 0.21\%$ and $r_2 = 0.745\%$. The equivalent of tables 1, 2, and 3 of the main text are given in tables A.1, A.2, and A.3. It is clear that the steady-state stability test in table A.2 is weaker than in the model presented in the main text (the stability hypothesis can only be rejected at the 12% significance level). This lesser degree of instability is probably due to the wider interval $[r_1, r_2]$. The steady-state location tests presented in table A.3 are also less conclusive. This may be due to the asymmetry of the interval $[r_1, r_2]$ around the average value of x, which in turn may be a result of asymmetry in the dynamics' nonlinear structure.

TABLE A.1

ESTIMATION OF A SETAR(2; 12, 12, 12) MODEL WITH $r_1 = 0.21\%$ AND $r_2 = 0.745\%$: U.S. INDUSTRIAL PRODUCTION, 1947:01-1988:04

	Equation $i = 1$	Equation $i = 2$	Equation $i = 3$
Estimated coefficients:			
ϕ_{i0}	0.002	- 0.003	0.0003
ϕ_{i1}	0.325	1.419	0.483
ϕ_{i2}	0.229	0.089	0.017
ϕ_{i3}	- 0.140	0.300	0.082
ϕ_{i4}	0.099	- 0.113	0.035
ϕ_{i5}	0.091	- 0.224	- 0.020
ϕ_{i6}	- 0.155	0.020	- 0.022
ϕ_{i7}	0.089	- 0.027	0.060
ϕ_{i8}	- 0.060	0.021	0.021
ϕ_{i9}	0.025	0.050	- 0.098
ϕ_{i10}	0.057	- 0.050	- 0.043
ϕ_{i11}	0.048	0.170	0.148
ϕ_{i12}	- 0.263	- 0.178	- 0.276
Steady-state of the estimated equation:			
x_i^*	0.30%	0.57%	0.05%
Sum of estimated lag-coefficients:			
$\sum_{j=1}^{k} \phi_{ij}$	0.345	1.465	0.369

TABLE A.2
TEST OF STEADY-STATE STABILITY[a]:

Number of observations used	Sum of lag-coefficients	Value of test-statistic	P-value
140	1.465	1.200	0.115

[a]Second equation of the SETAR(2; 12, 12, 12) model with threshold values: $r_1 = 0.21\%$ and $r_2 = 0.745\%$.

TABLE A.3
TESTS OF STEADY-STATE LOCATION[a]

Steady state	Estimate	Hypothesis	Value of test-statistic	P-value
(1) x_1*	0.30%	H_0: $x_1^* < r_1$	0.477	0.317
(2.1) x_2*	0.57%	H_0: $x_2^* < r_1$	1.409	0.079
(2.2) x_2*	0.57%	H_0: $x_2^* > r_2$	0.670	0.251
(3) x_3*	0.05%	H_0: $x_3^* > r_2$	3.856	0.0001

[a]Using a SETAR(2; 12, 12, 12) model with threshold values: $r_1 = 0.21\%$ and $r_2 = 0.745\%$.

REFERENCES

BROCK, W. A. AND C. L. SAYERS (1988): "Is the Business Cycle Characterized by Deterministic Chaos?" *Journal of Monetary Economics*, 22, 71-90.

DIAMOND, P. AND D. FUDENBERG (1987): "Rational Expectations Business Cycles in Search Equilibrium," *Journal of Political Economy*, forthcoming.

FRISCH, R. (1933): "Propagation Problems and Impulse Problems in Dynamic Economics," in *Economic Essays in Honor of Gustav Cassel*. London: George Allen and Unwin.

GOLDBERG, S. (1958): *Difference Equations*, New York: John Wiley.

GOODWIN, R. M. (1951): "The Non-Linear Accelerator and the Persistence of Business Cycles," *Econometrica*, 19, 1-17.

GUCKENHEIMER, J., AND P. HOLMES (1983): *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York-Berlin-Heidelberg: Springer-Verlag.

HAMMOUR, M. L. (1989a): "Increasing Returns and Endogenous Business Cycles," essay 1 in this thesis.

——— (1989b): "Indeterminacy and Instability in Macro Models with External Effects," essay 2 in this thesis.

HARVEY, A. C. (1981): *Time Series Models*. Oxford: Philip Allan Publishers.

HICKS, J. R. (1950): *A Contribution to the Theory of the Trade Cycle*. Oxford: Clarendon Press.

HOWITT, P. AND R. P. McAFEE (1988): "Animal Spirits," mimeo, University of Western Ontario.

MURPHY, K. M., A. SHLEIFER AND R. W. VISHNY (1988): "Increasing Returns, Durables and Economic Fluctuations," mimeo, University of Chicago.

SLUTZKY, E. (1937): "The Summation of Random Causes as the Source of Cyclic Processes," *Econometrica*, 5, 105-146.

STOCK J. H. AND M. W. WATSON (1988): "Variable Trends in Economic Time Series," *Journal of Economic Perspectives*, 2, 147-74.

TONG, H. (1983): *Threshold Models in Non-Linear Time Series Analysis*, New York-Berlin-Heidelberg-Tokyo: Springer-Verlag.