Discussion Paper No. 1079

Toward an Economic Theory of Pattern Formation

By Kimininori Matsuyama*

December 1993

*Department of Economics, Northwestern University, Evanston IL 60208, U.S.A. This note is written as a comment on the paper by Paul Krugman, "Complexity and Emergent Structure in the International Economy," given at the Conference, "New Directions in Trade Theory," the University of Michigan, October 29-30, 1993.

1. <u>Introduction</u>.

The real world is full of puzzles. As a kid growing up in downtown Tokyo, I always wondered why commercial districts surrounding different commuter rail terminal stations, like Ueno, Shibuya, Shinjuku, and Ikebukuro, all connected via the Circle Line, have such distinctive characters. Some sections of the city seemed to attract young people, while others appeal to older generations. Akihabara is full of discount electronics shops and Kanda full of bookstores. In school, I learned about how different industries are located in different parts of the country, and became aware of wide variations in income and growth across regions and wondered why. Booms and recessions are another phenomena that puzzled me (and they still do). In high school days, I found business cycles utterly incomprehensible; it seemed to me that any coming year is no different from the year that just went by; it is simply another round in the seemingly eternal motion of the earth around the sun. I remember challenging my classmates to explain to me how mere interactions among human beings could make one year more prosperous than others. In college, I majored in international relations and naturally began to wonder what generates the world system of rich and poor countries.

I thought that studying "modern economics," as the mainstream economics in North America is typically called in Japan, would help me understand all these puzzles; why are the economies are so full of patterns and variations, both spatial and temporal; and how do they emerge as the outcome of interactions among millions of people, or as the outcome of the "invisible hand." The result was a disappointment. In macroeconomics, they taught me that an increase in aggregate demand leads to a boom through multiplier effects, and yet they never explained to me why there is an increase in aggregate demand in the first place. In international economics, they taught me that a capital abundant country becomes an exporter of capital intensive goods, and yet they never told me why certain countries become more capital abundant in the first place. In development economics, they taught me that countries with high saving rates and education levels grow faster, and yet never told me why there are variations in saving rates and education across countries. In short, they responded to my question of why there are variations in some variables, by merely introducing variations in others. I thought that they were evading my question completely and felt cheated. My disappointment was so deep that, at one point, I seriously thought about pursuing Marxian economics or radical economics instead, not because I liked their philosophy, but because I thought that it would offer better chance of having my question answered, or at least taken seriously.

Now, after nearly fifteen years of additional knowledge, I feel a lot more positive about modern economics. Over the years, I have come across many, albeit isolated, attempts to explain spatial and temporal patterns and variations in economic activities. The most notable example is

Thomas Schelling's (1978) "Micromotives and Macrobehavior." The literature on endogenous business cycles, surveyed by Guesnerie and Woodford (1992), is another. In international economics, Krugman's work on the role of history in the determination of trade patterns (1990; Part II) and on economic geography (1991) has been a source of inspiration.

Meantime, I have also learned about several parallel movements across a wide range of natural sciences. They come under a variety of different names and buzzwords, such as "complexity," "positive feedback," "self-organization," "synergetics," "symmetry breaking," etc. But they are all trying to explain pattern formation in nature. Cosmologists wonder about why the gaps in Saturn's rings are spaced the way they are; why are galaxies spiral?; why is the matter in the universe distributed in clusters rather than evenly spread? Earth scientists use thermal convection models to explain formations of cloud streets, jet streams, ocean currents, as well as the earth's magnetic field and continental drifts. Material scientists study phase transitions, trying to understand how molecules aligned themselves when the materials are cooled down below the critical temperature. Molecular biologists ask how life began in the primordial soup of amino acids and developmental biologists, following the visionary work of D'Arcy Thompson and Alan Turing, attempt to explain how living organisms acquire beautiful forms through processes of cell division and morphogenesis, or how tigers and zebras got their stripes. Although the specific questions asked vary widely, scientists have come up with a coherent set of ideas in their efforts to explain the formation of patterns in nature.

In my opinion, economists ought to be more interested in explaining the formation of spatial and temporal patterns in economic activities, instead of merely assuming their existence. And I believe that they can learn a lot from the new developments in natural sciences. I am thus delighted to have the opportunity to read this paper by Paul Krugman, in which he points out some resemblances between his work on economic geography and a new branch of science called "complexity." Krugman argues that the science of complexity might be important for economists not so much because it offers a set of sophisticated mathematical models that they could appropriate, but rather because it offers "a useful organizing principle for thinking about some important issues in economics." I heartily agree: it seems to me that many of the important insights of the science of pattern formation or emerging structure, to use Krugman's terminology, is model-independent. Krugman wisely avoids technical details and mainly focuses on concepts in his paper. Following his strategy, I will also focus on the conceptual and methodological issues, in particular, those that I think deserve more attention than Krugman gave in his paper.

2. The Role of Complementarity in Pattern Formation.

The key concept for understanding pattern formation and emerging structure is the notion of complementarity. Broadly speaking, complementarities are said to exist when two phenomena reinforce each other. For example, if expansion of industry A leads to expansion of industry B, which in turn leads to the further expansion of A, then the two industries are complementary to each other. Or, if the arrival of a new plant in a town makes it a desirable location for other plants, then there are complementarities in the location decisions. The presence of such complementarity has profound implications for the stability of the system. If a change in a certain activity is initiated by a small external shock, this leads to a similar change in complementary activities and starts a cumulative process of mutual interaction in which the change in one activity is continuously supported by the reaction of the others in a circular manner. Many writers in the past, such as Hicks, Kaldor, Kalecki, Myrdal, and Nurkse, have stressed that cumulative processes of this kind should be an essential element in explaining business cycles, underdevelopment, and regional inequalities. On the other hand, the standard neoclassical model, with its efficient resource allocation, tends to emphasize the self-adjusting nature of market forces. As different activities compete for scarce resources, expansion of one comes only at the expense of others, which tend to dampen any perturbation to the system. The formation of pattern occurs when complementarities help the system to break away from the stabilizing forces of resource constraints.

To illustrate this idea, imagine an economy with two symmetric regions; each is endowed with an equal amount of an immobile factor, say land. In addition, there is a single mobile resource, say capital, whose allocation is measured along the horizontal axis in Figure 1. Production requires the use of both capital and land. Two curves show the rate of return to capital in the two regions for given regional allocation of capital. To the extent that capital has to compete for use of land, the rate of return declines as more capital is allocated in each region. When capital is divided evenly, the rate of return is equalized across the regions, and this is an equilibrium allocation. Furthermore, there is an obvious sense in which this equilibrium allocation of fifty-fifty is stable: if capital is allocated unevenly, the region with less capital offers a higher rate of return and hence attracts capital from the other region, and the resulting capital flow would restore the equilibrium allocation. To use Krugman's terminology, "centrifugal forces" of the resource constraint prevent one region from attracting more capital than the other.

Now let us introduce some agglomeration economies in capital allocation, or "centripetal forces" as Krugman might call them. Exactly why such agglomeration economies arise is not important for the moment. What is important is that sufficiently strong agglomeration economies would create a range in which the rate of return increases with capital, as in Figure 2. As "centripetal forces" dominates "centrifugal forces", there will be complementarity in the allocation of capital. The model now has two stable equilibria and both imply uneven spatial allocations of capital, say, twenty-eighty and eighty-twenty. This is exactly what we wanted: starting with the symmetric physical environment, we end up reaching asymmetric spatial outcomes. The symmetric allocation of capital is still an equilibrium, but it is unstable. If a slightly more than half of capital is allocated in one region, then the rate of return would be higher there and hence more capital would flow from the other region. Once this process get started, it would feed on itself, and the allocation of capital would be further away from the fifty-fifty division.

I am sure that many of you must have seen diagrams like Figure 2, and this kind of argument should be quite familiar. But I suspect that some of its implications may be less familiar. Let me discuss a couple of them.

3. The Logic of Symmetry Breaking and Multiple Equilibria

If you want to explain the formation of pattern, or emerging structure, your model must possess a high degree of symmetry or uniformity in the physical environment. The goal is to explain spatial and temporal variations of economic activities as an outcome of internal mechanisms of the market system. The simple model above illustrated how it can be done. With complementarities, asymmetry in endogenous variables can be generated in spite of perfect symmetry in exogenous variables.

So, here's a puzzle. How could symmetric causes can lead to asymmetric effects? Shouldn't the symmetry in the cause be somehow reflected in the effect as well? The answer is yes. The symmetry in the cause must be translated into the symmetry of the set of possible effects. That is, if the twenty-eighty allocation is a possible outcome, so is the eighty-twenty allocation. This means, however, that each possible outcome does not have to be symmetric if there is more than one possible outcome. Of course, there is always an equilibrium, which possesses the same degree of symmetry as the environment. And yet, such a symmetric equilibrium may lack stability and hence does not represent the most plausible outcome. In fact, it is the loss of stability of the symmetric equilibrium that generates the symmetric set of asymmetric equilibria, which leads to the pattern formation or emerging structure. Scientists call this fundamental process of pattern formation, "broken symmetry," or "symmetry breaking."

There is nothing particularly deep about this. It is a simple mathematical statement: a solution to a symmetric equation need not be symmetric; only the set of solutions has to be symmetric. Once pointed out, this may seem quite obvious to you. And yet, economic journals are littered with articles, in which the authors claim that the equilibrium is symmetric because of the symmetric structure of the model. Even when they are aware of asymmetric equilibria, they often suggest that the symmetric equilibrium is a natural one to consider.

One immediate corollary is this. If you insist that your model has to have a unique equilibrium, there is no hope of explaining pattern formation: with the conditions that ensure the uniqueness of equilibrium, the symmetry of the cause must carry over to the effect. Thus, *if you want to develop a theory of pattern formation, you have to accept multiplicity of equilibria.*

Many economists, however, seem to feel religious about uniqueness of equilibrium. So, what would they do if they are asked to explain something like: the capital/land ratio in region one is four times as high as in region two? They would, of course, argue that the technology is better in region one, as a justification for drawing two curves representing the rates of return so as to fit the observed variations, as in Figure 3. Belief in the virtue of uniqueness forces them to explain one variation by introducing another.

If you think that I am critical of models with unique equilibrium, you are wrong. To the contrary, I think they are extremely powerful tools for understanding how variations in one set of variables would translate into another set of variables: that is what comparative static analysis is all about. The problem is that economists are trained to perform comparative statics, which have been so successful; some economists even start thinking that performing comparative statics is the sole purpose of building a model, which makes them hostile to any model with multiple equilibria. But, if so, they have forgotten about a whole set of important issues, such as why the real world looks the way it does, which cannot be answered without building a model with multiple equilibria.

International trade is one of many fields that suffer from this tendency. The neoclassical theory of trade relies entirely on the differences in technology and in resource endowment, which are themselves unexplained. As Kaldor (1970) argued so convincingly, however, very little of these exogenous variables is truly external to the economic system. In particular, for many items in resource endowment, such as physical capital, human capital, R&D, it is difficult to separate cause from effect. Unfortunately, most people in this field do not seem to feel any obligation to explain them. I have been to so many seminars, where speakers defend the assumption of technological differences across countries by merely saying, "after all, that's what international economics is about."

This attitude, I think, is responsible for making the neoclassical theory look so sterile and unattractive to economic historians, and political scientists (as well as some college students like me fifteen years ago), who often turn to more radical and structuralist approaches for a guide to understand the emergence of the world trade system. Krugman's recent research on economic geography is a powerful statement that the mainstream economics, based on the methodological individualism, can also address these issues of critical importance. I am hoping that Krugman's work, or something along this line, would help us reestablish communication channels to a much broader range of social scientists.

Here's some good news. By accepting multiplicity, we do not need to give up models with the unique equilibrium, or the knowledge we have accumulated through the use of such models. Remember that a model of multiple equilibria does not mean that many things happen at the same time; it only means that many things are possible. In the real world, of course, something definite has to happen. Once a particular stable asymmetric equilibrium is selected, a particular pattern has been formed, or a particular structure has emerged, then economists can treat it as given and go about estimating parameters under the maintained hypothesis that something like Figure 3 is a good approximation (in the neighborhood of the state that is actually observed, of course) and then performing comparative statics. So, there is no reason for them to be hostile to models with multiple equilibria.

As a great mathematical physicist of this century, Hermann Weyl, once wrote in his famous lecture, *Symmetry*, "The truth as we see it today is this: The laws of nature do not determine uniquely the one world that actually exists."

4. The Logic of Sudden Change

One possible objection for models with multiple equilibria would be that they cannot yield useful predictions. My response to this objection is that the very fact that multiple equilibria exist for certain parameter values and not for others is itself a prediction of great significance. Unfortunately, most economists who work with multiple equilibria tend to impose restrictions to ensure the multiplicity. Similarly, those economists who like to perform comparative statics tend to impose at the outset the restrictions that ensure uniqueness. What I like most about Krugman's modelling strategies is that he always introduces two competing forces in such a way that multiple equilibria exist only in a subset of the parameter space. By changing those parameters, we could thus tell when and how multiple asymmetric equilibria can be generated: in other words, we could "witness" the emergence of structure in his models.

Going back to Figures 1 and 2, what differentiates these two cases is the relative magnitude of the centripetal to centrifugal forces, which may in turn be determined by some parameters of the model, say, λ. Figure 4 schematically traces out how stable equilibrium allocations change as one changes the parameter λ continuously. It is precisely by their choice of such a parameter that theorists demonstrate their ingenuity. The trick is to find a parameter that is not so obvious, in order to maintain a reasonable distance between assumptions and conclusions. Such a parameter must be measurable independently of the phenomena that need to be explained. Of course, each theorist may have his or her own favorite: In Krugman's case, it is the transportation cost. Whatever one's favorite parameter may be, there is always a critical value of the parameter, λ_c, at which there is an abrupt change in the nature of the model. In Figure 4, if $\lambda < \lambda_c$, the equilibrium is unique and stable; no spatial asymmetry appears. On the other hand, $\lambda > \lambda_c$ implies the presence of two stable asymmetric equilibria. So, as λ exceeds its critical value, patterns have been formed. It turns out that this is a fairly generic feature of models in which two competing forces are at work.

Some physical analogies may be apt here. If you heat water at atmospheric pressure, nothing drastic happens until the temperature reaches 100°C, when it suddenly started boiling. If you lower the temperature, water suddenly turns into ice at 0°C. Such abrupt changes in the phase of matter are the results of a competition between the attractive intermolecular forces, which tend to order the system, and the individual kinetic energies of the molecules, which have the opposite effect. In a liquid phase, the random motion of water molecules is too violent for intermolecular interaction to hold them in place. Increasing the pressure (which favors the intermolecular forces) or decreasing the temperature (which reduces the random motion of molecules) would eventually tip the balances between the two. Once the critical point is crossed, the attractive forces will be strong enough to keep molecules firmly in place and the crystal structure of ice will be developed. There are many similar examples in nature. At room temperature, iron can be magnetized, but when it is heated above the Curie point, it loses its ferromagnetic property. Another example would be superconductivity, in which electronic resistance disappears discontinuously at the critical temperature.

Figure 4 also shows that the rate at which a stable asymmetric equilibrium responds to a change in λ is arbitrarily large just to the right of the critical level. This has to be the case as long as the model depends continuously on λ. This is another paradox of pattern formation due to symmetry breaking. *At the onset of pattern formation, a change in the environment can be amplified to equilibrium allocation at an arbitrarily large rate.*

The moral is clear and should be familiar: the last straw could break camel's back. Yet, the economist schooled in Alfred Marshall's dictum, *Natura non facit saltum*, may endlessly search for a big change in the environment when he tries to explain a large change in a certain economic variable. But, his effort might very well turn out be futile. Or when somebody comes up with a hypothesis, such as the big change in X was caused by a change in Y, he may criticize and reject it on the ground that the change in Y observed is too small to be responsible for the change in X. But the logic of pattern formation suggests that his criticism may be unwarranted.

One natural question, which must already have crossed your mind, is how a particular stable asymmetric equilibrium actually gets chosen out of many, as the parameter crosses its critical level. One possibility is to argue that a purely random event will dictate the selection. Maybe, a particularly aggressive individual happens to live in one region, or an engineer in one region comes up with an idea of a new production method. Or maybe, a crazy politician happens to live in another. Or maybe, (using Krugman's favorite example) a teenaged girl named Catherine Evans, who lived near Dalton, Georgia, happened to make a bedspread as a wedding gift. Or maybe, George Bailey of Bedford Falls (played by James Stewart in Frank Capra's *It's a Wonderful Life*) was never granted his wish. And these individuals, like those described in detail by Hughes' *The Vital Few* (1986), can act as a catalyst in the process of regional pattern formation.

My favorite theory of selection, which differs from the random selection theory more in perspective than in substance, is based on imperfection of our knowledge. Perfect symmetry exists only in mathematics; in the real world, things are not perfectly symmetric; there are some minor differences across regions that even a most acute observer could not detect. Introducing such an asymmetry in the structure of the model could change Figure 4 into something like Figure 5. Now, as you see, as λ gets large, more and more capital gets allocated to region one. The outcome is thus deterministic, but *even a small asymmetry in the environment can lead to a large asymmetry in equilibrium*. Since our knowledge is less than perfect, for all practical purposes, we could not predict it. Nor would simple linear regressions be able to discover any cause responsible for such a large-scale asymmetry in effect.

According to the random selection theory, God plays fair dice. According to my favorite theory, the dice are loaded, but we just don't know how. I like the second better because it suggests the presence of some biases in selection. Although it would be impossible to predict in advance, historians, with their advantage of hindsight, may be able to "postdict" how the selection occurs. With a large number of careful case studies, they may be able to find some "patterns" in the process of pattern formation.

I do not know which theory of equilibrium selection you like better. But both theories suggest the significant role of historical research. Economics of pattern formation may help us bridge the gap between economic theory and economic history.

5. Wider Implications

So, why should we care about developing a theory of pattern formation and emerging structure? The real world is full of patterns and variations; we, economists, should be able to explain the observed patterns in economic activities, or at least try to explain them. To my mind, this is a reason sufficiently good enough to support the kind of the research pursued by Krugman. But, some practical economists, or the National Science Foundation, may find such a "curiosity-driven" research program unsatisfactory. To these people, I would argue that trying to explain the formation of pattern would also help us address old questions from a new angle. For instance, think about the North-South problem. Recently, many economists have tried to answer the question posed by Lucas (1990), "Why Doesn't Capital Flow from Rich to Poor Countries?" They typically approach this problem by taking the presence of rich and poor countries for granted, and come up with fairly convoluted solutions. To my mind, the right approach is to step back and ask ourselves, "Why Are There Rich and Poor Countries?" I do not attempt to answer this question here, but it seems to me that, if you come up with an answer to the second question, it would automatically suggest a solution to the first question. Another example would be the urbanization problem in developing countries. The literature is dominated by the Harris-Todaro model, which assumes that the economy already has two different areas, one urban and the other rural. But how could we evaluate problems of urbanization without understanding how one area becomes urbanized and the other remains rural? Or think about the potential impacts of economic integration in the context of Europe and NAFTA? Most mainstream economists are preoccupied with numerical exercises, which generally show that the concerns raised by structuralist economists are unfounded. But in the models they use, the parameters are chosen to match the observed variations of the industrial structure across the economies. Krugman's regional models, capable of explaining the regional patterns of industrial localizations as an internal mechanism of the market system, suggest that the possibility of such an structuralist nightmare cannot be dismissed so easily.

A theory of pattern formation may have much wider implications even beyond international and interregional trade and the problem of underdevelopment. For example, think of the occupational distribution across ethnic groups. Taking the stereotypes, the Greek run restaurants, the Chinese laundries, and the Korean groceries. If you apply the standard model of occupational choices, like the Roy model, such occupational patterns force us to conclude that there are large intrinsic differences in skills across ethnic groups. A theory of pattern formation would suggest that there are some complementarities in the processes of skill acquisition, limited within an ethnic group, so that some small differences in skills or some random events happen to end up sorting different ethnic groups into different occupations. Or think of the problem of comparative economic systems. It is now widely acknowledged that there are large systemic differences across national economies, particularly in labor markets. For example, it has been pointed out that the seniority system, corporate union, lifetime employment, firm-specific on-the-job training, frequent rotations in job assignments, and bonus payments, are far more prevalent in Japan than in the United States. There are very strong temptations to attribute these observations to the differences in regulations (which is implicit in the Structural Impediment Initiatives talks) or even to cultural differences (which lead to revisionist sentiments). However, a theory of pattern formation would suggest an alternative view. Due to some institutional complementarities across different employment practices and across different firms, either most firms in a country would adopt these practices, or else, very few firms would adopt them. Only some historical accidents happen to make them more prevalent in Japan than in the United States.

Finally, the conceptual framework discussed here may also help us understand better the emergence of social customs as well as cultural differences across societies. If this turns out to be the case, it could provide a useful organizing principle for a wide range of social sciences, from anthropology to economics to sociology, just as the new science of pattern formation has developed as an interdisciplinary venture across a variety of natural sciences.

References:

Guesnerie, Roger, and Michael Woodford, "Endogenous Fluctuations," Advances in Economic Theory: Sixth World Congress, vol. II, edited by J.J. Laffont, (Cambridge University Press: New York), 1992.

Hughes, Jonathan, The Vital Few, Expanded Edition, (Oxford University Press: New York), 1986.

Kaldor, Nocholas, "The Case for Regional Policies," Scottish Journal of Political Economy, November 1970, 337-348.

Krugman, Paul, Rethinking International Trade, (MIT Press: Cambridge), 1990.

Krugman, Paul, Geography and Trade (MIT Press: Cambridge), 1991.

Lucas, Robert E. Jr., "Why Doesn't Capital Flow from Rich to Poor Countries?" American Economic Review 80 (May 1990): 92-96.

Schelling, Thomas C., Micromotives and Macrobehavior, (W. W. Norton Company: New York), 1978.

Weyl, Hermann, Symmetry (Princeton University Press: Princeton), 1969.

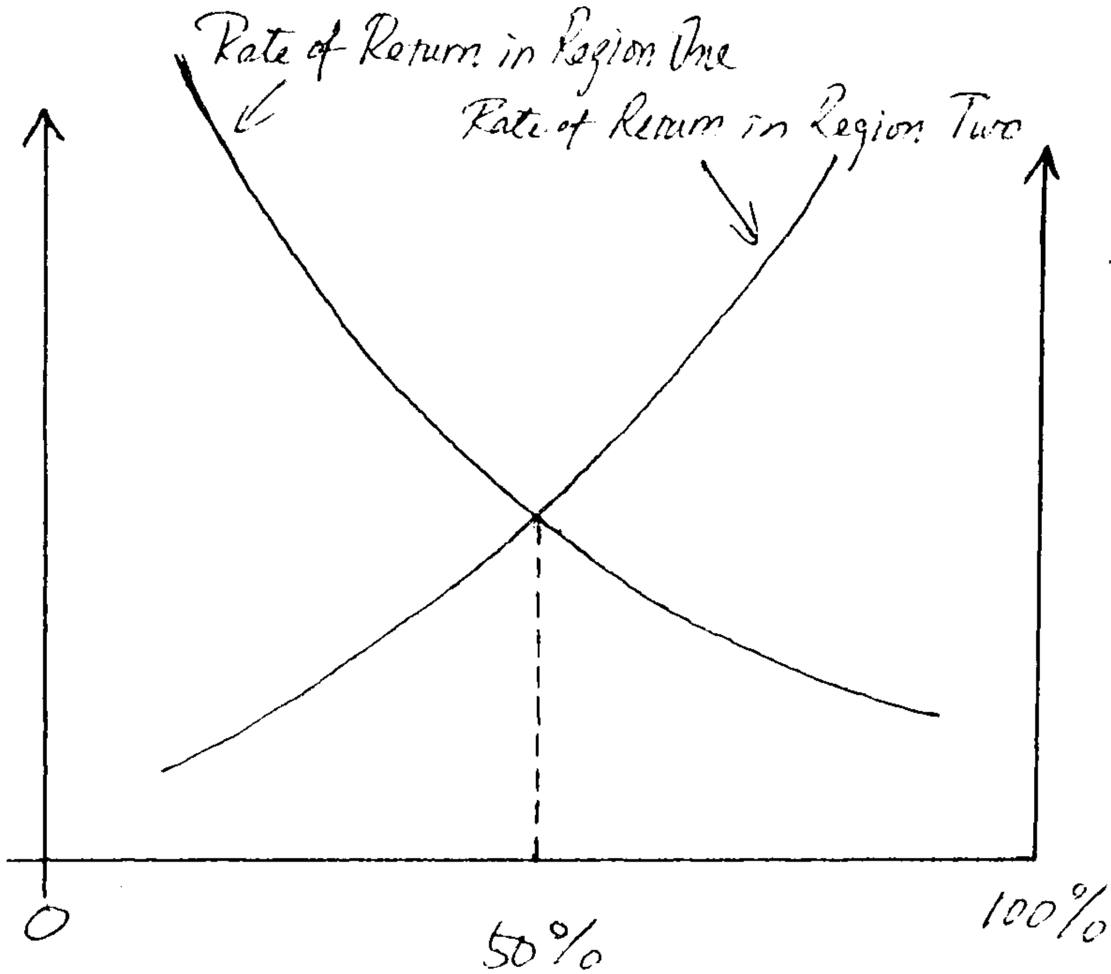
Rate of Return in Region One
Rate of Return in Region Two
0
50%
100%

Figure 1

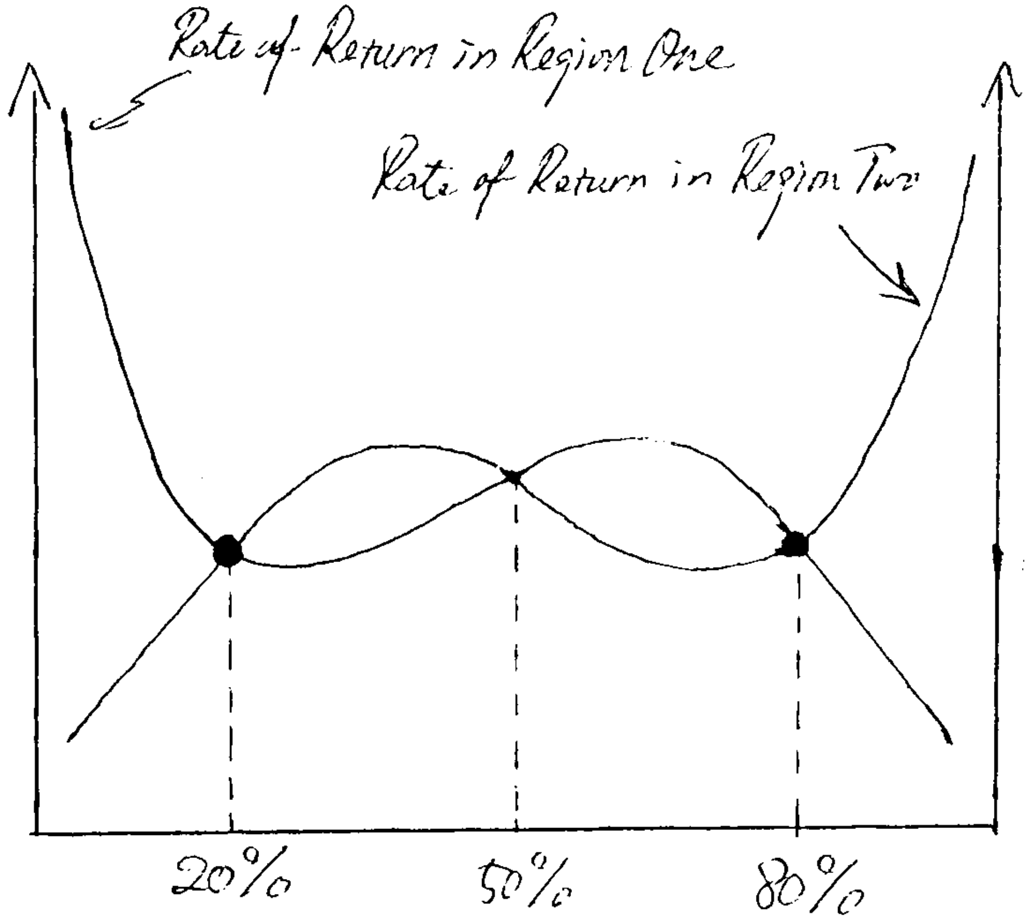
Rate of Return in Region One
Rate of Return in Region Two
20%
50%
80%

Figure 2

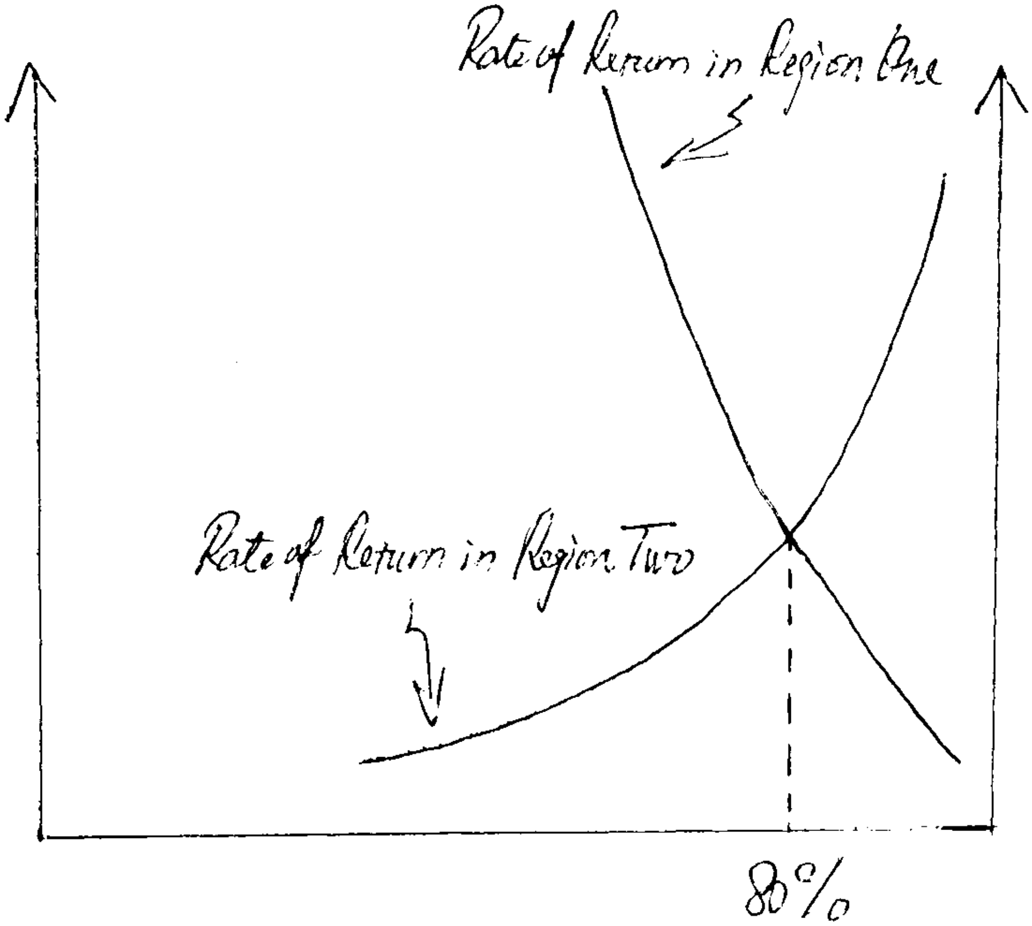

Figure 3

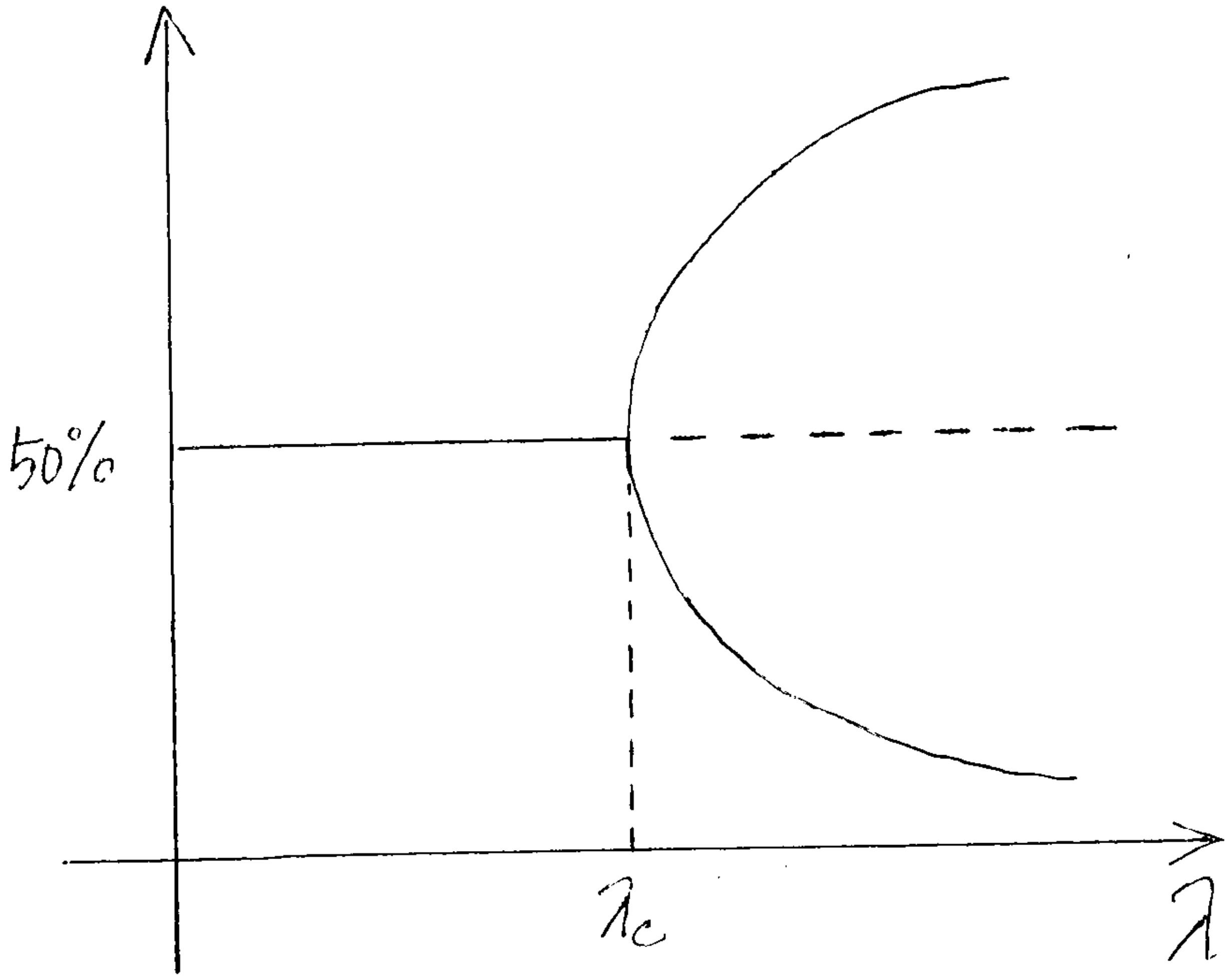

Figure 4.

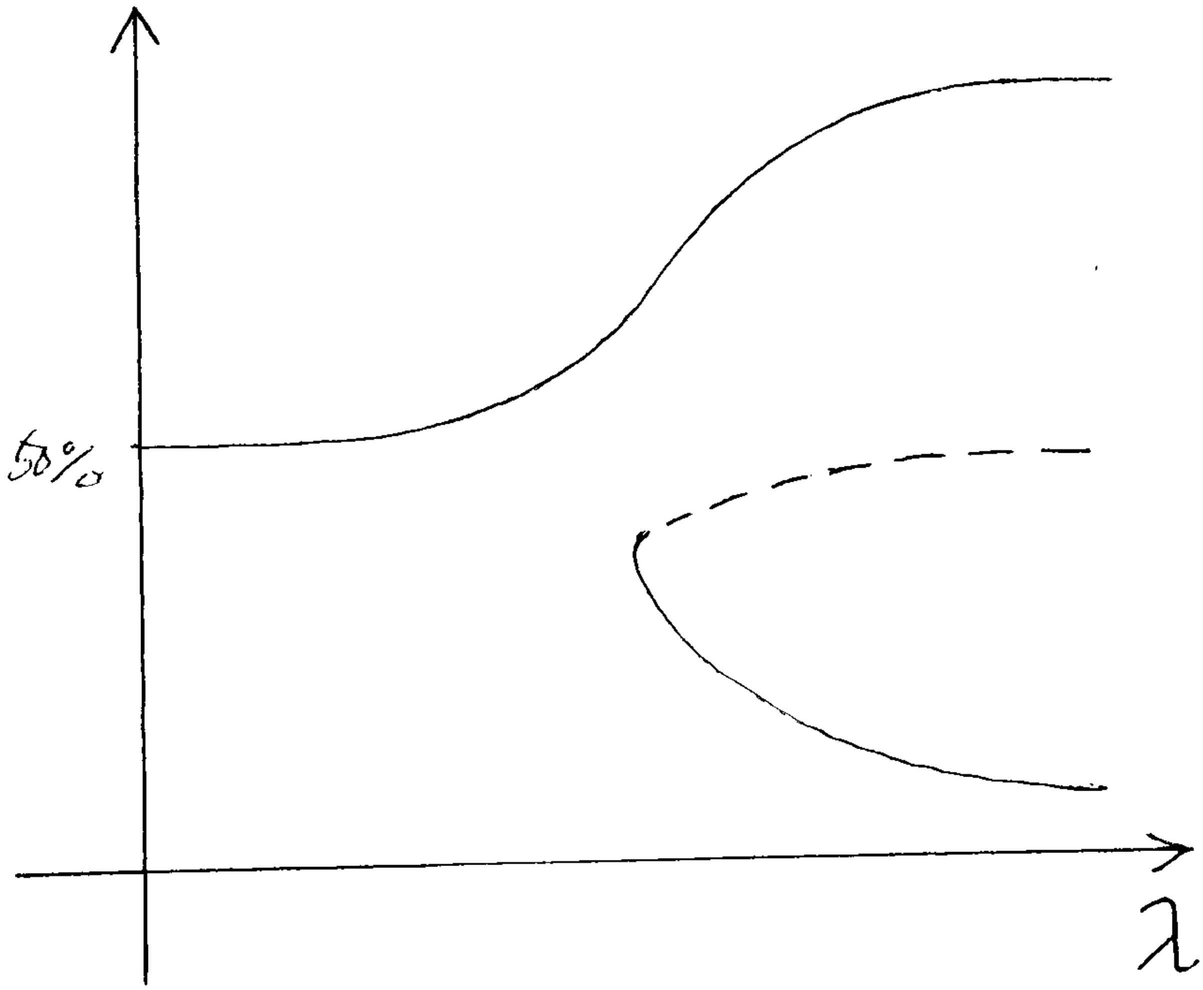

Figure 5