Albert Hirschman
The Strategy of Economic Development
Yale University Press 1957

CHAPTER 4

Unbalanced Growth: An Espousal

Is Balance in Supply Required?

IN THE LAST CHAPTER we criticized the idea that development must take place simultaneously in many activities to provide the element of "mutual support" that alone will make it possible to clear the market of the newly produced goods. Having discarded this "pure" theory of balanced growth we must still consider a far less rigorous version, one that insists that if growth is not to be stunted the various sectors of an economy will have to grow jointly in some (not necessarily identical) proportion; no sector should get too far out of line, not because of demand but because of supply or "structural" considerations. For instance, if secondary industry grows, the food and raw material input needed by the workers and the machines will go up; if some of these requirements are imported, then an increase in exports is necessary, etc., etc.

In this form, the balanced growth theory is essentially an exercise in retrospective comparative statics. If we look at an economy that has experienced growth at two different points in time, we will of course find that a great many parts of it have pushed ahead: industry and agriculture, capital goods and consumer goods industries, cars on the road and highway mileage—each at its own average annual rate of increase. But surely the individual components of the economy will not actually have grown at these rates throughout the period under review. Just as on the demand side the market can absorb "unbalanced" advances in output because of cost-reducing innovations, new products, and import substitution, so we can have isolated forward thrusts on the supply side as inputs are redistributed among users through price changes, and at the cost of some temporary shortages and disequilibria in the balance of payments or elsewhere. In fact, development has of course proceeded in this way, with growth being

communicated from the leading sectors of the economy to the followers, from one industry to another, from one firm to another. In other words, the balanced growth that is revealed by the two still photographs taken at two different points in time is the end result of a series of uneven advances of one sector followed by the catching-up of other sectors. If the catching-up overreaches its goal, as it often does, then the stage is set for further advances elsewhere. The advantage of this kind of seesaw advance over "balanced growth," where every activity expands perfectly in step with every other, is that it leaves considerable scope to *induced* investment decisions and therefore economizes our principal scarce resource, namely, genuine decision-making.

Classical economics, while not taking so positive a view of the imbalances of the growth process, at least was never particularly concerned about them because it relied on prices to signal, and on the profit motive to eliminate rapidly and reliably, any structural disequilibria that might arise in the course of growth. The critics of classical economics, on the other hand, have always pointed to cases in which these "market forces" would not act with adequate strength and speed. Having thus convinced themselves that the adjustment mechanism is beset with virtually insuperable obstacles, some of the critics naturally enough took the defeatist view that growth has to be balanced from the start or cannot take place at all.

This counsel of perfection is not only impractical but also uneconomical. We need not sacrifice the valuable development mechanisms brought into play by unbalanced growth, especially if we go beyond the overly narrow view of the adjustment process that has long dominated economic literature.

Tradition seems to require that economists argue forever about the question whether, in any disequilibrium situation, *market forces acting alone* are likely to restore equilibrium. Now this is certainly an interesting question. But as social scientists we surely must address ourselves also to the broader question: is the disequilibrium situation likely to be corrected at all, by market or nonmarket forces, or by both acting jointly? *It is our contention that nonmarket forces are not necessarily less "automatic" than market forces.* Certainly the almost monotonous regularity with which interventionist economists have come forward—and with which authorities have acted—when

the market forces did not adequately perform their task testifies to the fact that we do not have to rely exclusively on price signals and profit-maximizers to save us from trouble.[1]

The case of unbalanced growth provides a good illustration. When supply difficulties arise in the course of uneven progress in sectors—such as education and public utilities—where private enterprise is not operating, strong pressures are felt by public authorities to "do something"; and since the desire for political survival is at least as strong a motive force as the desire to realize a profit, we may ordinarily expect some corrective action to be taken.[1a]

There is no implication here that any disequilibrium whatsoever will be resolved by some combination of market and nonmarket forces. But if a community cannot generate the "induced" decisions and actions needed to deal with the supply disequilibria that arise in the course of uneven growth, then I can see little reason for believing that it will be able to take the set of "autonomous" decisions required by balanced growth. In other words, if the adjustment mechanism breaks down altogether, this is a sign that the community rejects economic growth as an overriding objective.

The inclusion of probable reactions of nonmarket forces not only serves to make economic analysis more realistic. It also protects us

1. Some traditional equilibrium mechanisms were unable to dispense entirely with help from agents outside the market. Thus, the restoration of balance-of-payments equilibrium and the damping of the business cycle was, for a long time, made to depend on correct manipulation by the central bank of the rate of interest, in reaction to developing disequilibria. But this role of the central banker has usually been rationalized as an exception to the rule; and in the minds of many economists, the central banker became a sort of honorary member of the market forces.

1a. Sectoral imbalances have of course been a conspicuous feature of Russian economic development. The resulting difficulties have been described in Soviet literature as "nonantagonistic contradictions" which are not only admitted to exist but apparently considered to perform a useful signaling and corrective function: "The characteristic trait of our difficulties and contradictions consists precisely in that they themselves indicate to us the basis and the means for their solution." V. Kozlovskii, *Antagonisticheskie i neantagonisticheskie protivorechiia* (Moscow, Moskovskii Rabochii, 1954), p. 70. These "nonantagonistic" contradictions which are successfully overcome by administrative action of the Communist party and the government are then opposed to the "antagonistic" contradictions which are said to afflict capitalism and which can be resolved only by revolution.

against a fallacious chain of reasoning that is fairly common in development economics and of which the doctrine of balanced growth is itself an excellent illustration. In this reasoning, one first selects some objective of economic policy that seems desirable enough; then one proves that the objective cannot be attained through the operation of market forces; and one concludes that state action surely will bring the objective about. But this conclusion is clearly a non sequitur. The fact that private entrepreneurs will be unable or unwilling to do certain jobs which we would like to see done does not in itself ensure that the government can handle them. We must examine whether these jobs are likely to be performed satisfactorily by public authorities, which function after all in the same society as the entrepreneurs.[2]

Development as a Chain of Disequilibria

As has been shown, the balanced growth theory results from comparing the initial point of underdevelopment equilibrium with another point at which development will practically have been accomplished. A certain impatience with the process that lies between these two points—i.e., for the process of development—is shown by the following quotation from a well-known article by Scitovsky:

> Profits are a sign of disequilibrium; and the magnitude of profits under free competition may be regarded as a rough index of the degree of disequilibrium. Profits in a freely competitive industry lead to investment in that industry; and the investment in turn tends to eliminate the profits that have called it forth. Thus far, then, investment tends to bring equilibrium nearer. The same investment, however, may raise . . . profits in other industries; and to this extent it leads away from equilibrium. . . . The profits of industry B created by the lower price for factor A, call for investment and expansion in industry B one result of which will be an increase in industry B's demand for industry A's product.

2. Much the same point is made forcefully by Bauer and Yamey with respect to governmental promotion of industrial enterprise: "A general lack of enterprise in a country does not in itself set up a presumption of such initiative in the public sector," *Underdeveloped Countries*, p. 161. However, I do not follow the authors in the conclusions which they draw for the role of governments in economic development. See Chs. 8 and 11.

> This in turn will give rise to profits and call for further investment and expansion in A; and equilibrium is reached only when successive doses of investment and expansion in the two industries have led to the simultaneous elimination of investment in both. It is only at this stage that . . . the amount of investment profitable in industry A is also the socially desirable amount. The amount is clearly greater than that which is profitable at the first stage before industry B has made its adjustment. We can conclude, therefore, that when an investment gives rise to pecuniary external economies, its private profitability understates its social desirability.[3]

To my mind, the first part of this passage is a most pertinent portrayal of how development is set and kept in motion, but Scitovsky, considering the proceedings he describes unnecessarily laborious, proposes to short-circuit them and to reach in a single jump a new point of equilibrium where the "elimination of investment" has been accomplished. But, actually, development is a lengthy process during which interaction of the kind described by Scitovsky takes place not only between two industries, but up and down and across the whole of an economy's input-output matrix, and for many decades. What point in such a virtually infinite sequence of repercussions are we supposed to shoot at? Which intermediate expansion stages ought we to skip, and which ordinarily successive stages ought we to combine? Some skipping or combining may be possible, but with no more than the modest objective of speeding up development here and there. In general, development policy must concern itself with the judicious setting up of the kind of sequences and repercussions so well described by Scitovsky, rather than with any attempt to suppress them. In other words, our aim must be to *keep alive* rather than to eliminate the disequilibria of which profits and losses are symptoms in a competitive economy. If the economy is to be kept moving ahead, the task of development policy is to maintain tensions, disproportions, and disequilibria. That nightmare of equilibrium economics, the endlessly spinning cobweb, is the *kind* of mechanism we must assiduously look for as an invaluable help in the development process.

Therefore, the sequence that "leads away from equilibrium" is pre-

3. "Two Concepts of External Economies," pp. 148–9.

cisely an ideal pattern of development from our point of view: for each move in the sequence is induced by a previous disequilibrium and in turn creates a new disequilibrium that requires a further move. This is achieved by the fact that the expansion of industry *A* leads to economies external to *A* but appropriable by *B*, while the consequent expansion of *B* brings with it economies external to *B* but subsequently internal to *A* (or *C* for that matter), and so on. At each step, an industry takes advantage of external economies created by previous expansion, and at the same time creates new external economies to be exploited by other operators.[4]

In Scitovsky's example, these external economies are essentially caused by production complementarities of one type or another, and we are thus returning to the complementarity effect of investment which was already invoked in Chapter 2 as a mechanism that would make investment decisions particularly easy or compelling. We were then speaking of the investment-promoting character of investment, not indirectly through additional savings out of the incomes created by investment, but through direct contact or "contagion."

Technical complementarity in the strict sense is usually defined as a situation where an increase in the output of commodity *A* lowers the marginal costs of producing commodity *B*. This will happen typically as a result of the following situations:

a. because *A* is an input of *B* and is produced under conditions of decreasing costs;

b. because *B* is an input of *A* and is itself produced under conditions of decreasing costs;

c. because *A* and *B* are joint products (or because *B* is a by-product of *A*) and are produced under decreasing costs.

Because situations such as these have long been familiar to economists, complementarity is usually associated with economies of scale.[5] But there is no need for so restrictive an interpretation. We can define complementarity as any situation where an increase in

4. Note that the private profitability falls short of the social desirability of any venture only when its "output" of external economies exceeds its "input" derived from other ventures.

5. W. Fellner, *Trends and Cycles in Economic Activity* (New York, 1956), pp. 199–200. N. S. Buchanan and H. S. Ellis, *Approaches to Economic Development* (New York, 1955), pp. 279–80.

the demand for commodity *A* and the consequent increase in its output call forth an increased demand for commodity *B* at its existing price. This happens not only when the connection between the two commodities is via the production process. The connection between *A* and *B* may also arise because the increased *use* of *A* leads to greater demand for *B*. We are not thinking here of situations where *A* and *B* *must* be employed jointly in fixed proportions. In this case it would not make much sense to say that demand for *A* and the subsequent increase in its output provide an incentive for the production of *B*, as it is rather the demand for the good or service into which *A* and *B* enter jointly which explains the demand for both products. This is the familiar case of derived demand. But there are many situations in the course of economic development where the increased availability of one commodity does not *compel* a *simultaneous* increase in supply of another commodity, but *induces slowly*, through a loose kind of complementarity in use, an upward shift in its demand schedule. The phenomenon has been described under the apt heading "entrained want"; [6] Veblen observed it long ago and effectively summed it up when he said that "invention is the mother of necessity" rather than vice versa.

An example of the rigid type of complementarity in use (best treated as derived demand) is cement and reinforcing steel rods in the construction, say, of downtown office buildings. Examples of the looser, "developmental" type of complementarity (entrained want) can be found in the way in which the existence of the new office buildings strengthens demand for a great variety of goods and services: from modern office furniture and equipment (still fairly rigid), to parking and restaurant facilities, stylish secretaries, and eventually perhaps to more office buildings as the demonstration effect goes to work on the tenants of the older buildings. Here again, failure to arrange for all of these complementary items from the start could be denounced

6. The term is used by H. G. Barnett in *Innovation: The Basis of Cultural Change* (New York, McGraw-Hill, 1953), pp. 148–51, with the exact meaning we have in mind here: "The fulfillment of one need establishes conditions out of which others emerge . . . In most instances it is impossible for people to foresee [these emergent wants] even if they try . . . Entrained wants are a consistent feature of motivational stresses for cultural change" (p. 148).

as "poor planning" which ought to be avoided by centralized decision-making. But, just as in the case quoted by Scitovsky, an attempt to telescope the whole process would be futile because of the virtually infinite number of complementarity repercussions, and because of the uncertainty about a good many of them; moreover, such an attempt would miss the point that the profitable opportunities that arise as a result of the initial development move constitute powerful and valuable levers for subsequent development which are to be carefully nursed, maintained at some optimum level, and if necessary created consciously rather than eliminated.[7]

The common feature of the various complementarity situations is that, as a result of the increase in the output of *A*, the profitability of the production of *B* is being increased because *B*'s marginal costs drop, or because its demand schedule shifts upward, or because both forces act jointly.

Put even more generally, complementarity means that increased production of *A* will lead to *pressure* for increasing the available supply of *B*. When *B* is a privately produced good or service, this pressure will lead to imports or larger domestic production of *B* because it will be in the *interest* of traders and producers of *B* to respond to the pressure. When *B* is not privately produced, the pressure does not transmute itself into pecuniary self-interest, and will take the form of political pressure for the provision of *B*. This is the case for such public services as law and order, education, satisfactory monetary and banking arrangements, highways, water, electric power, etc. Complementarity then manifests itself in the form of complaints about

7. This does not mean that when new buildings are put up one should refrain from planning for new parking facilities. Development itself constantly extends the range of complementarities that are rigidly compelled and necessarily simultaneous: the optional equipment of one period becomes the standard equipment of the next, as a result of social and cultural pressures and needs rather than because of purely technological factors. The process of turning loose complementarities into rigid ones is often called "integrated planning" which is then opposed to "improvisation." These terms, particularly dear to city planners, are quite misleading in their antagonism. "Integrated planning" takes care of *a few* of the *known* repercussions of a development move rather than letting them take care of themselves as best they can independently of that move. But it certainly can never hope to comprehend them all.

shortages, bottlenecks, and obstacles to development. Action in this case does not take place through the operation of the profit motive, but through group pressures on public authorities and agencies.

A Definition of Induced Investment

The complementarity effect provides us with a new concept of *induced* investment which is more meaningful for underdeveloped economies than the conventional one, i.e., investment that is directly related to past increases in output. For this conventional concept of induced investment has validity mainly for countries with a fully built-up industrial and agricultural structure where increases in demand lead to increases in capacity designed to keep marginal costs from entering the area in which they would begin to rise steeply. The required adjustments may cover many industries, but are ordinarily small in any one year in relation to existing capacity. The big dynamic changes in developed economies are expected to originate in "autonomous" investment.

This is not a realistic picture of the growth process in underdeveloped economies. Here an increase in the demand for beer, for example, may lead not only to the expansion of existing brewing capacity but, at a certain point, to the *start* of domestic production of bottles, of barley cultivation, and to a whole chain of similar repercussions. In other words, the investment that is induced by complementarity effects may help to bring about a real transformation of an underdeveloped economy.

One of the difficulties of the concept of induced investment in its traditional meaning is its precise delimitation. The reason for which investment is undertaken is not that demand has increased in the past, but that the experience of the past is taken as a guide to the future. In other words, investment is undertaken because for one reason or another the ensuing output is expected to find a market. But looked at in this way, all investment is obviously induced and the distinction between induced and autonomous investment becomes untenable or arbitrary.[8]

At first blush, it might seem that the same flaw, in an even more pronounced form, affects the distinction we have drawn. Is not every

8. Fellner, *Trends*, p. 319.

investment "induced" in the sense that it complements some other existing investment? With the generously wide definition of complementarity which we have given, cannot every step in the development of a country be considered as called forth by the preceding steps in a never-ending series of "inducements"? Have we then perhaps explained too much?

At this point we may, however, revert to our earlier discussion of external economies: it was then shown that new projects often appropriate external economies created by preceding ventures and create external economies that may be utilized by subsequent ones. Some projects create more external economies than they appropriate and therefore their private profitability falls short of their social desirability. It is therefore to be expected that the opposite situation can also be encountered—namely, ventures that have a large "input" of external economies and a much smaller "output." The projects thus favored represent the class of "easy-to-exploit" investment opportunities which always abound in newly developing economies.

We can then define our concept of induced investment by the provision that the projects that fall into this category must be *net beneficiaries* of external economies.

This definition makes induced investment look very much like the multiplier: each investment is conceived as inducing a series of subsequent investments and there is an element of convergence as the "output" of external economies diminishes at each step. This, however, does not necessarily mean that the investments themselves converge; there is no rigid connection between the size of an investment and its net "input" of external economies, although some association between these two magnitudes may be expected to exist.[9]

Theoretically, our definition of induced investment is, I believe, more satisfactory than the conventional one and it is far more relevant in the context of development problems. Nevertheless it is extremely difficult to give empirical content to the concept and we shall therefore not attempt to give our reasoning more rigor than it possesses; we shall continue to speak of investment inducing other investments and shall simply be aware that there are widely varying degrees of "inducements."

An ideal situation obtains when, as was pointed out in the last sec-

9. See Ch. 6, the concept of satellite industries.

tion, one disequilibrium calls forth a development move which in turn leads to a similar disequilibrium and so on ad infinitum. If such a chain of unbalanced growth sequences could be set up, the economic policy-makers could just watch the proceedings from the sidelines. It may be noticed that in this situation private profitability and social desirability are likely to coincide, not because of the absence of external economies, but because "input" and "output" of external economies are the same for each successive venture.

In practice, growth sequences are likely to exhibit tendencies toward convergence or potentialities of divergence, and development policy is largely concerned with the prevention of too rapid convergence and with the promotion of the possibilities of divergence.

One more point. The induced investment defined thus far is a gross quantity. In the previous chapter, we made much of the point that development brings with it external diseconomies as well as economies. The external diseconomies brought into the world by new investments refer primarily to the damage done to existing industrial or handicraft establishments by the introduction of modern methods and products. It must be granted, therefore, that new investments will hold back reinvestment in these establishments while leading to complementary capital formation elsewhere in the economy. The effect is, however, quite asymmetrical, as the greatest damage that new investment will cause to pre-existing equipment consists in failure to maintain and replace that equipment. Thus, inasmuch as the external diseconomies of new investments result in negative investment, this destructive effect is likely to be spread over several years; whereas the positive effect of the external economies leads at once to a demand for the total capital requirements of whatever ventures are going to be "induced." Because of this asymmetry, the investment-reducing effect of new investments resulting from competition and substitution effects seems unlikely to match the investment-creating effects of complementarity except where competitive industries are strong and complementarity effects rather weak. This latter situation may be characteristic of the textile industry and may account for the fact that in several underdeveloped countries the setting up of this industry has failed to provide the necessary spark for further development.

Some Related Points of View

The way in which investment leads to other investment through complementarities and external economies is an invaluable "aid" to development that must be consciously utilized in the course of the development process. It puts special pressure behind a whole group of investment decisions and augments thereby that scarce and non-economizable resource of underdeveloped countries, the ability to make new investment decisions.

The manner in which an investment project affects the availability of this resource is for us the principal measure of its contribution to further development. A development strategy that stems from this approach is outlined in the next chapters. Before closing the present one we shall refer briefly to some development theories that are related to the point of view presented here.

One of the principal characteristics of our approach has been the direct connection we have established between the investment of one period and that of the next. The complementarity effect "calls forth" new investment; to the extent that savings are determined by this process, they play a perfectly passive role. This situation is very similar to the one recently described by Domar in his analysis of a growth model elaborated thirty years ago by the Russian economist Feldman. The essence of this model is the division of total investment into investment designed to expand the output of consumers' goods on the one hand, and of producers' goods, on the other. Domar shows that once this division is made "the propensity to save has no life of its own so to speak and is completely determined by the relative productive capacities of the two categories." [10] The limit to investment in the Feldman model as interpreted by Domar is not the ability or propensity to save, but the productive capacity of the investment goods sector. Although the model is admittedly unrealistic, especially for an open economy, it is interesting as an attempt to build a sequence where investment of one period is directly related to the investment of prior periods without the intermediary of the savings ratio.

The rather tormented discussion around an article on investment criteria by Galenson and Leibenstein has yielded a similar line of

10. *Essays*, p. 236.

CHAPTER 5

Investment Choices and Strategies

Efficient Sequences versus Investment Criteria

WE CAN now begin to consider one of the most crucial problems in development theory and policy: that of investment choices.

Development requires the undertaking of a series of projects producing favorable effects on the flow of income, in a wide variety of fields: public administration, education, health, transportation, power, agriculture, industry, urban development, etc. The limitation of resources, be they savings available for investment or our "ability to invest," compels a choice among these projects. In traditional economics, the market performs this function by equating the productivities of the various projects at the margin. It is recognized, however, that in any economy a substantial proportion of funds must be devoted to projects (in education, health, some public utilities, etc.) whose output has no readily assigned or fully recoverable market value. Moreover, underdeveloped economies tend to exhibit certain systematic discrepancies between private costs and social costs, and in such cases reliance on the market would lead to misallocation of resources.[1]

These considerations and the practical needs of development planners have led to the elaboration of *investment criteria.* The problem that has been discussed in this connection can be formulated as follows: given a limited amount of investment resources and a series of proposed investment projects whose total cost exceeds the available resources, how do we pick out the projects that will make the greatest contribution relative to their cost? In answering this question, econ-

1. There are at least three important areas in which such systematic discrepancies are apt to occur: the wage rate (because of disguised unemployment), the exchange rate (because of overvaluation of the currency), and the interest rate (because of rationing of loan funds on the part of the banks). See J. Tinbergen, *The Design of Development* (Baltimore, 1958), pp. 39 ff.

omists have ordinarily interpreted "contribution" as *direct contribution to output* once the project has been completed. This is only natural if growth is visualized as depending exclusively on aggregate output and income which, via the propensity to save, secretes the means for further growth. On these premises, the measurement of what has been called the "social marginal productivity" (SMP) of different projects—essentially a more or less sophisticated benefit-cost ratio—becomes the instrument that should in theory permit us to rank different projects in the order of their expected contribution to output and therefore to further growth.[2]

Recently, a far more elaborate concept has been proposed by Leibenstein: In addition to the output stream, investment criteria ought to take account also of the differential effects of the proposed ventures on the supply of entrepreneurship and of savings, on consumption habits, population increases, and a variety of other factors affecting further growth.[3] Leibenstein admits that a criterion embodying all these repercussions (in addition to SMP proper) would be of unusually difficult application.[4] In practice, his criticism seems likely to result in an agnostic "it all depends" attitude since it seriously impairs the usefulness of the SMP criterion without replacing it by a manageable new instrument.

In attempting a different approach, we shall first draw a distinction between substitution choices and postponement choices. Consider any choice between project *A* and project *B:* If the decision favors *A*, this may mean either that *B* is *discarded permanently* or that it is *postponed.* In the former case, the choice is between technical substitutes such as alternative means of providing a city with power or water supply. Many important choices are of this kind. They relate to the best means of attaining a given end or to the best design of a project whose output itself is needed beyond question. In deciding such choices, the usual investment criteria retain considerable usefulness.

2. A. E. Kahn, "Investment Criteria in Development," *Quarterly Journal of Economics, 55* (Feb. 1951), 38–61; H. B. Chenery, "The Application of Investment Criteria," *Quarterly Journal of Economics, 57* (Feb. 1953), 76–96; J. Ahumada, "Preparación y evaluación de proyectos de desarrollo económico," *El trimestre económico, 22* (July–Sept. 1955), 265–96.

3. Leibenstein, *Economic Backwardness and Economic Growth,* ch. 15.

4. Ibid., p. 268.

Nevertheless, we feel that in underdeveloped countries additional considerations must be introduced and we will do so in Chapter 8.

Let us suppose for the time being that all substitution choices have been made and that we have before us a series of useful projects which are ideally designed to accomplish their respective purposes. In this situation, we are only faced with postponement choices.[5] We no longer choose *A* instead of *B;* rather, we choose the sequence *AB* instead of the sequence *BA*. What is the possible rationale for such a choice? If we suppose that our goal is to have both *A* and *B*, but that "now" we can undertake only either *A* or *B*, leaving *B* or *A*, respectively, for "later," then it is clear that the only conceivable reason for preferring *AB* to *BA* is that *B* will be possible sooner once *A* is in place than vice versa. In other words, our choice depends entirely on the pressure that the existence of *A* exerts toward the coming into existence of *B* as compared to the corresponding pressure that would emanate from *B* toward *A*. Once the problem is formulated in this way it becomes quite clear that the comparative productivity of *A* and *B* which will both have to be undertaken is likely to be a rather minor factor in the decision assigning the priority.

Although our reasoning has been drastically simplified, it takes hold of an important aspect of the development problem. Essential tasks always abound in underdeveloped countries since backwardness has so many different interrelated facets. From this interrelatedness we do not draw the balanced growth conclusion that a simultaneous attack is essential. But what might be called a sequential or chain solution is indeed required. In other words, isolated progress in one area is possible, but only for a limited period; if it is not to be choked off, it

5. In an earlier paper, "Economics and Investment Planning: Reflections Based on Experience in Colombia" in *Investment Criteria and Economic Growth,* ed. M. F. Millikan (Cambridge, Mass., M.I.T., 1955, multilithed), I argued essentially that economists ought to confine themselves to the making of substitution choices. I still believe that the most urgent task of development planners usually consists in arriving at correct substitution choices; but I realize now that postponement choices cannot be evaded. They must be made at two different stages of the process of development planning: first before it is decided in which sector or sectors substitution choices are to be studied, for the decision seriously to study alternative means of fulfilling a given need usually already implies a decision to give priority to this need; and secondly, after substitution choices have been completed in several different sectors.

must be followed by progress elsewhere. Therefore to compare the productivity increases that result from two projects in, e.g., education and transportation, is an insoluble problem not only in practice but conceptually. Such comparisons must be made on the *ceteris paribus* assumption that progress is being achieved in only one of the areas; and on this assumption the longer-term productivity of both undertakings is simply *zero* since the improved transportation facilities will serve little purpose and will fast deteriorate if education is not also improved in due course and vice versa. Therefore, the question of priority must be resolved on the basis of a comparative appraisal of the strength with which progress in one of these areas will induce progress in the other.[6] In these basic types of development decisions, it is therefore not sufficient to supplement, qualify, and otherwise refine the usual investment criteria. We must evolve entirely new aids to thought and action in this largely uncharted territory of efficient sequences and optimal development strategies.

There is no doubt that the task that we have set ourselves is extremely complex. Let us suppose that we know which are the *n* steps that need to be taken to, say, double a country's per capita income. Then there exist in principle *n!* possible sequential arrangements of these *n* steps! Of course, there can be no question of neatly deducing, through a series of syllogisms, *the* most efficient sequence. Rather, we will strive to "suboptimize" [7] and to develop a few guideposts, principles, and illustrative models.

To begin with, there was a great deal of exaggeration in our statement that there exist *n!* sequences in which the *n* steps may be undertaken. Many sequences are unavoidably "one-way" for purely tech-

6. It may be objected that indivisibility could not be such as to prevent us from investing our resources partly in education and partly in transportation. However, the point we are making does not depend on indivisibility in the sense of "lumpiness." Let us assume that we have identified *n* essential and interrelated projects, costing 200 million dollars, but that we have only 100 million dollars at hand. Suppose that out of the *n* projects we can put together various collections of $m<n$ projects costing 100 million dollars. Then again the criterion for picking any particular collection of *m* projects would be the strength with which their execution would induce the remaining projects. Thus indivisibility is assumed only in the trivial sense that some projects will necessarily be undertaken ahead of others.

7. Charles Hitch, "Sub-optimization in Operations Problems," *Journal of the Operations Research Society of America, 1* (May 1953), 87–99.

nical reasons (a road must be built before it can be paved); one also feels that other one-way sequences are imposed not because they are technically determined but because they are necessary if development is to be properly planned, i.e., is to proceed in an "orderly" fashion. But here there may be some doubt as to how far it is advisable to go. Observation tells us that rapid growth of countries, cities, industries, and individual firms hardly ever proceeds in a completely orderly fashion, but that an excess of disorderliness may exert an inhibiting and demoralizing influence on further growth. Can we then perhaps define an optimum degree of orderliness in development? To illustrate this problem, let *A*, *B*, *C*, and *D* in Figure 2 represent a group of development steps we wish to take and that ought to be

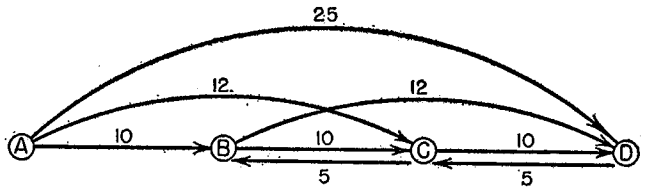

Fig. 2. The optimum disorderliness model

taken in this order if ideal "orderliness" is to be achieved. Let us also suppose that step *A* ***must*** be realized before *B*, *C*, or *D* can possibly be undertaken, but that with *A* accomplished the sequence is no longer imposed. In the absence of limiting factors, the sequence *ABCD* would be chosen because it provides the smoothest transition from state *A* to state *ABCD*. But we now introduce a limited resource, such as decision-making or organizational ability, or simply time, and assume that different amounts of this resource are spent in going from one point to another. We want to minimize the use of this resource. If, say, ten units of this resource are spent in going from *A* to *B*, from *B* to *C*, and from *C* to *D*, then it is natural to think that to go from *A* directly to *C* will take a somewhat larger (say 12 units) and from *A* to *D* perhaps a much larger amount (say 25 units), because of the absence of the intermediate preparatory stages. On the other hand, less than ten units (say 5) should be needed to "fill in" *B* or *C* after *C* or *D*, respectively, because once the later steps have been realized the lack of the intermediary ones makes itself felt in so pressing a manner that the decision to undertake them requires far smaller quantities of the

scarce ability or time than when they represented genuine forward steps.

If we apply the foregoing illustrative figures, then the expenditure of our scarce resource that is involved in the various possible sequences is as follows:

A to *B* to *C* to *D*	30 (10 + 10 + 10)
A jump to *C* then fill in *B*, then *D*	27 (12 + 5 + 10)
A to *B* then jump to *D*, then fill in *C*	27 (10 + 12 + 5)
A jump to *D*, then fill in *B* and *C*	35 (25 + 5 + 5)

In this example the figures have been selected so as to show that a limited amount of "putting the cart before the horse" may be efficient as compared to both maximum orderliness and maximum disorderliness.

It may be helpful to attempt a translation of this model into more familiar terminology. Let us assume two ventures, *m* and *n*, which require equal amounts of capital and have a yield of 10% and 8% respectively. At the beginning of period 1 the interest rate stands at 9%, hence only venture *m* is undertaken. At the beginning of period 2, with venture *m* in existence, the expected yield of venture *n* has risen to 10% and is now also launched. But we are free to suppose that, if *n* were undertaken first, *m* would be urgently required and that its expected yield would rise to 14% at the beginning of period 2. In this eventuality, investors would maximize income by selecting in period 1 the investment with the lower yield! Besides they would do everything to rush *m* to completion. Such strange results are avoided in traditional theory by the implicit assumption which we chose to discard here, that the profitability of different ventures is invariant with respect to the order in which they are undertaken.

The preceding examples are highly artificial as they imply that development proceeds along a single path. Nevertheless, they embody a number of concepts that are recurring throughout this essay: the difference between "permissive" and "compulsive" sequences, the possible rationality of violating "first things first" norms and the fact that the difficulty of taking a development decision is not necessarily proportional to the amount of capital it requires.

A more complex and perhaps more realistic model would be to consider development as the putting together of a jigsaw puzzle. The fitting in of individual pieces would represent the taking of discrete

development steps. The problem would again be to minimize the time needed to put the puzzle together. The total time is of course equal to the sum of the time periods spent on fitting in the individual pieces, and the time needed for each piece could be made to depend inversely on the number of contacts with adjacent pieces already in place: with each piece surrounded by several neighbors, the larger the number of neighbors in place, the less time it will take to find and fit into its proper place the common neighbor of these neighbors. Each fitting is more or less "induced," depending on the ease or difficulty with which it may be made.[8] An efficient sequence for putting the puzzle together could be found by trial and error once we have information about the varying amounts of time needed for fitting in individual pieces. For instance, if the time needed fell rapidly toward zero the larger the number of neighbors already in place, then the efficient sequence would turn out to be completely different from the one that would be optimal if the increase in the facility with which pieces may be fitted in were subject to decreasing returns as the number of neighbors increased.

Up to this point, we have considered that the difficulty of taking any development steps (i.e., the fitting in of the individual jigsaw pieces) depends exclusively on the number of neighbors already in place. We can bring our model one step closer to reality by supposing that the taking of the different steps varies in intrinsic difficulty *besides* being affected by the number of neighbors. If this is the case, then the putting together of the puzzle becomes far more determinate than before: for now we would aim at surrounding by "neighbors" those pieces that are intrinsically most difficult to fit in, securing thereby far

8. In the usual jigsaw puzzle the task of fitting in a piece also becomes progressively easier as the game progresses and the number of remaining loose pieces declines. Although this feature of a jigsaw puzzle could be related to the "take-off" concept and to Simon's learning model (cf. Ch. 2), it is rather a disturbing element from the point of view of the problem which we wish to illustrate at this juncture. To eliminate it, we may imagine that the jigsaw puzzle goes on forever: only a limited number of loose pieces can be chosen from at any point of time, but as soon as one piece is fitted, a new one is mixed in among the loose pieces on the table. Such a representation of our model is consistent with the view that the growth process is an infinite one, but that at any one point of time only a limited number of steps-to-be-taken is within the horizon of the decision-makers.

greater economies of effort than if we surrounded those pieces that are intrinsically of average or less-than-average difficulty.

These fanciful digressions may illustrate the kind of models in terms of which a general theory of "efficient sequences" might be built. I doubt, however, that it is useful to go very far in this direction. Our short discussion had primarily the purpose:

1. to make the concept of efficient sequence a little more palpable; and
2. to show that efficient sequences will necessarily vary widely from region to region and from country to country depending on the location and stubbornness of the principal development difficulties.

Social Overhead Capital versus Directly Productive Activities

Definitions and biases. The distinction between Social Overhead Capital (SOC) and Directly Productive Activities (DPA) is a recent one. Like all such classifications it must be judged not by its logic, which is far from compelling, but by its theoretical and practical usefulness, which has been considerable. SOC is usually defined as comprising those basic services without which primary, secondary, and tertiary productive activities cannot function. In its wider sense, it includes all public services from law and order through education and public health to transportation, communications, power and water supply, as well as such agricultural overhead capital as irrigation and drainage systems. The hard core of the concept can probably be restricted to transportation and power. Thus limited, SOC can be operationally defined as comprising those activities for the financing of which the International Bank for Reconstruction and Development shows a pronounced preference, just as the behavioral sciences have been said to comprise all those endeavors which manage to obtain financial support from the Ford Foundation. The conditions for including an activity under the category of SOC are probably at least the following three:

1. The services provided by the activity facilitate, or are in some sense basic to, the carrying on of a great variety of economic activities.
2. The services are provided in practically all countries by public agencies or by private agencies subject to some public control: they

are provided free of charge or at rates regulated by public agencies.

3. The services cannot be imported.

The difference between the wide and the narrow meaning of SOC depends on whether one adds a fourth condition, namely:

4. The investment needed to provide the services is characterized by "lumpiness" (technical indivisibilities) as well as by a high capital-output ratio (provided the output is at all measurable).

This last condition clearly focuses attention away from, say, health and education, toward port installations, highways, hydroelectric projects, etc.

Statistical and historical research has shown the importance of SOC in the total investment picture as well as the large share of foreign investment that went into SOC, particularly railroads, during the nineteenth and early twentieth centuries.[9] As a result, economists, and particularly the "developers" among them, have become acutely SOC-conscious. It is widely assumed that enlarged availabilities of electric power and of transportation facilities are essential preconditions for economic development practically everywhere. Here, at least, we have a field where economists have given full recognition to the principle of "efficient sequence." Investment in SOC is advocated not because of its direct effect on final output, but because it permits and, in fact, invites DPA to come in.

The trouble with investment in SOC—or is it its strength?—is that it is impervious to the investment criteria that have been devised to introduce some rationality into development plans. The computation of capital-output ratios often presents almost insuperable statistical difficulties (as in the case of highways) and is moreover considered to be misleading anyway because of the igniting effect SOC investment is expected to have on DPA. As a result, SOC investment is largely a matter of faith in the development potential of a country or region.

The fact that there is so little possibility of evaluating objectively how much investment in SOC is really indicated in any given situation should give us pause. Such a situation implies at least the possibility of wasteful mistakes.

9. Cairncross, "The Place of Capital in Economic Progress"; and Nurkse, "International Investment Today in the Light of Nineteenth Century Experience," *Economic Journal, 64* (Dec. 1954), 744–58; and *Problems of Capital Formation,* pp. 152–4, and articles by Carter Goodrich there cited.

The absence of ex ante criteria is compounded by the weakness of sanctions when mistakes have actually been made. Underutilized port installations, highways, and even power plants do not present nearly the same administrative and public relations problem as a factory that is idle or suffers losses because of insufficient demand.

Perhaps it is this absence of criteria and of sanctions that has endeared SOC so much to the developers. Development planning is a risky business and there is naturally an attraction in undertaking ventures that cannot be proven wrong before they are started and that are unlikely ever to become obvious failures.

It must be conceded that, to some extent, investment in SOC is "safer" than investment in DPA, not only on this account but also in a real sense: it is diversified investment in the general growth of the economy rather than in the growth of one specific activity. This is the case, for instance, of improvements in a country's *principal* port, of modernization of an *integrated* railroad system, and of additions to the capacity of an *interconnected* electric power production and distribution system. But many SOC investments do not represent similarly diversified risks, and are rather narrowly tied to the anticipated rising fortunes of one city, one valley, or one traffic route; in such cases, it is questionable whether SOC investment is less risky than, for example, investment in an industry whose products might have a nationwide market.

There is another important reason for which the importance of SOC investment may have been overstressed in recent years. In the countries outside the Soviet orbit, directly productive activities in industry, agriculture, and commerce are generally in the hands of individuals or private firms. Therefore, development programs, although often claiming to lay down comprehensive patterns of resource use for the future, are concerned primarily with the allocation of public investment funds among those activities that are considered to be the responsibility of public agencies. The provision of transportation and communications facilities, the production and distribution of electric power, the construction of irrigation and drainage systems are now widely agreed to be appropriate fields of governmental economic activity in addition to the more traditional ones of law and order, defense, education, health, etc. Since economic planners are then spending most of their time on SOC projects, it is only natural that they should claim for them over-

riding and fundamental importance. This in itself would be innocuous enough were it not for the fact that a combination of taboos, opposing interests, and self-restraint makes it difficult for public investment to enter the DPA sectors. Therefore SOC investment is not only being overadvertised; it also risks being overdone since alternative and possibly more desirable uses of public funds are simply not within the horizon of the planners.

Development via shortage and via excess capacity of SOC. There can be no question whatever that SOC investment is "essential" for economic development. The sizable percentage of total investment occupied by SOC investment in all countries testifies to this fact. But all we know from such statistics is that SOC investment is a most important ingredient of economic development. *They cannot tell us to what extent SOC investment leads or follows DPA investment,* and this is the question we are interested in.

There is no simple answer. Either sequence is conceivable and we must look toward economic analysis and history for indications of the advisability of one or the other under given conditions.

Some SOC investment is required as a prerequisite of DPA investment. Access to an area by sea, road, rail, or air is indispensable before other economic activities can unfold there. But within rather wide limits, the relationship between SOC and DPA is not technologically determined. Within these limits, the cost of producing any given output of DPA will be the higher, the more inadequate the SOC of the economy. This situation is shown in Figure 3. Availability (and cost) of SOC is measured on the horizontal axis. *Total* cost of producing DPA output (including depreciation charges on DPA investment) is shown on the vertical axis. It is assumed that SOC investments do not enter DPA cost calculations. The *a* curve indicates the cost of producing a given full-capacity output of DPA from a given DPA investment as a function of the availability of SOC. The *b*, *c*, and *d* curves show this cost for successively higher amounts of DPA output from successively higher DPA investments.

The slope of the curves is easily explained. Starting at the far right, SOC is plentiful and DPA costs are low; additional SOC hardly makes for any further decline in DPA costs. As we move left, costs for any given DPA output rise slowly at first, then more rapidly; the cost

curves eventually become vertical as there is a minimum SOC which is indispensably required for producing any given DPA output.

The cost curves are superficially reminiscent of the familiar isoquants which show output as a function of two inputs such as labor

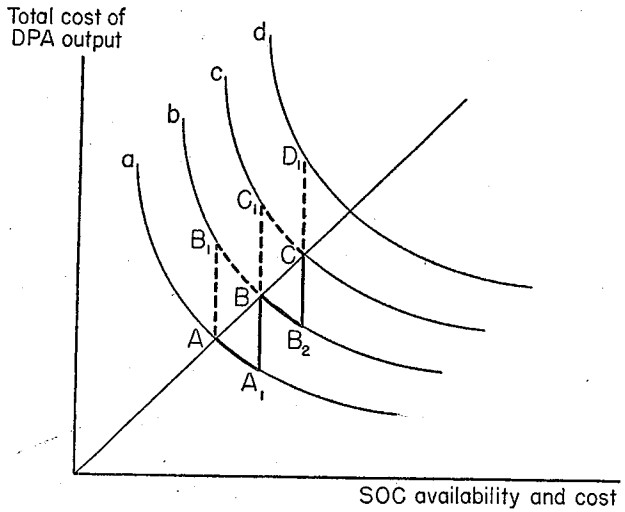

Fig. 3. Balanced and unbalanced growth of DPA and SOC

and capital. Actually the situation is quite different. Our cost curves do not show how a given output can be produced by increasing one input and decreasing another: they are strictly two-dimensional since they reflect the variations in the cost of a given output as only one element, namely the availability of SOC, is being altered. And this element is not even an input in the usual sense, since it does not constitute an "internalized" cost for the DPA producers.

From the point of view of the economy as a whole, the objective is to obtain increasing outputs of DPA at minimum costs in terms of resources devoted to both DPA and SOC. On each curve, the point the *sum* of whose coordinates is smallest is therefore the most desirable one. We have drawn the curves in such a way that the 45° line through the origin connects the optimal points of the different curves. This line expresses the ideal of balanced growth of DPA and SOC: a bit of each at each step no doubt would result in the greatest economy of the country's resources.

But it is one of the paradoxes of development that poor countries cannot always afford to be economical. Our principal assumption throughout this essay is that the real scarcity in underdeveloped countries is not the resources themselves but the ability to bring them into play. We shall now apply this notion by stipulating:

1. that SOC and DPA cannot be expanded at one and the same time; and

2. that preference should go to that sequence of expansion steps which maximizes "induced" decision-making.

As a result of the first condition, we can immediately visualize two principal types of sequence: one that starts expansion through increases in the supply of SOC—shown in Figure 3 by the fat line connecting points AA_1BB_2C—and one where the initial expansionary step is always taken by DPA, indicated in the figure by the dotted line AB_1BC_1C. The first sequence may be labeled "*development via excess capacity*" (of SOC) and the latter "*development via shortage*" (of SOC).

According to the second condition, our preference should go to the sequence that shows the greatest promise of being vigorously self-propelling. It is difficult to tell from our figure which one is likely to be superior on this ground. If we start by expanding SOC (sequence AA_1BB_2C), existing DPA production becomes less costly and an increase in DPA investment therefore may well get under way, depending on the response of entrepreneurs to increased profits. If, on the other hand, expansion of DPA is undertaken first, DPA production costs are likely to rise substantially and DPA producers will realize the possibility of making considerable economies through the installation of larger SOC facilities. As a result, pressures for such an increase are likely to come into play and will thus induce the next step in this sequence.

Both sequences therefore set up incentives and pressures, and an evaluation of their respective "efficiency" depends on the strength of entrepreneurial motivations on the one hand and on the response to public pressure of the authorities responsible for SOC on the other.

It may be asked how the sequence could start at all by expanding DPA, in view of the rise in costs that such an expansion, unaccompanied by a concurrent enlargement of SOC, would entail. The answer is that, even at B_1, DPA may still be profitable. Moreover, we have made

no institutional assumptions about the way in which either DPA or SOC is organized. In other words, we have not debarred ourselves from undertaking public investment in DPA should we find that the more efficient development sequence is set in motion in this way. Also, if that is our finding, it may represent an important argument for protection and other forms of subsidy to private industry. We have here a new type of infant industry argument: in a situation where SOC is not plentiful it may be more efficient to protect, subsidize, provide special finance for, or to undertake directly investment in DPA than to stimulate DPA indirectly through investment in SOC.

A combination of the two types of sequences is also conceivable. Suppose that investment decisions of both SOC and DPA operators need the special push of SOC shortage or SOC excess capacity, respectively, but that once the push is felt, *both overreact* to it and thereby bring about a pressure situation of the opposite type: then we would obtain a third type of sequence which in the figure would weave around the balanced growth line. For instance, the reaction of the SOC operators to the shortage situation prevailing at B_1 would bring DPA costs down all the way to B_2, whereupon DPA operators would overreact to the excess capacity of SOC by jumping to line *d*, thereby creating a new SOC shortage at D_1, etc. Such a mixed "staircase" sequence that continually overshoots "balance" and thereby sets up alternate types of investment pressures has interesting points of similarity with the cobweb model. From the point of view of development behavior and strategy, however, the two pure sequences seem more meaningful and will now be considered in greater detail.

Suitability of the various sequences. The principal characteristic of the two varieties of unbalanced growth which we have described is that they yield an extra dividend of "induced," "easy-to-take," or "compelled" decisions resulting in additional investment and output. Excess capacity of SOC, "building ahead of demand," is expected to create this demand by making a country, region, or city attractive to DPA investors. If, on the other hand, DPA is allowed or is made to run ahead of SOC, strong pressures are set up for the provision of SOC in a subsequent period. Development via shortage is an instance of the "disorderly," "compulsive" sequence discussed earlier in this chapter. Development is speeded up because an intermediate stage

that is being jumped over can be filled in with comparative ease as pressures and needs arise from the already realized stages. The absence of the intermediate stage is now felt as a shortage and the decisions to remedy it are more readily taken than before the shortage arose. For instance, fiscal measures and utility rate changes which are needed to secure the funds for expansion and whose adoption was out of the question prior to the shortage are suddenly accepted as inescapable.

Thus balanced growth of SOC and DPA is not only unattainable in underdeveloped countries; it may not even be a desirable policy be-

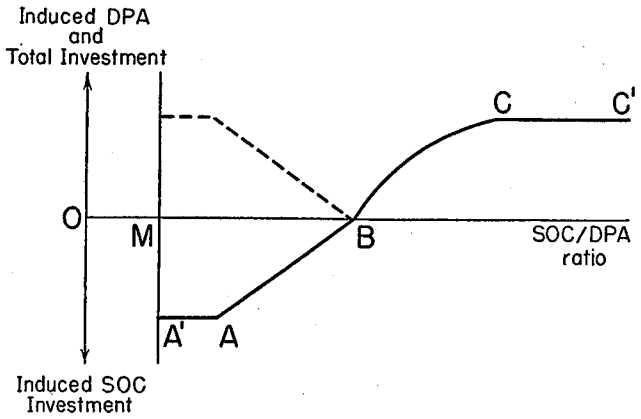

Fig. 4a. Investment induced by SOC shortage or excess

cause it does not set up the incentives and pressures that make for this "dividend" of induced investment decisions.

Let us try to be more explicit. Suppose that, during a first period, a certain number of SOC and DPA investment decisions have been made; they have resulted in some ratio of SOC to DPA output. We are now interested in analyzing the *induced* investment decisions that result from this ratio in a second period. This problem is illustrated in Figures 4a and 4b. Along the horizontal axis we plot the value of the SOC/DPA ratio. We plot induced DPA investment on the vertical axis in the upward direction and induced SOC investment along the same axis but in the downward direction. Total induced investment is shown by a dotted curve or line in the upward direction to the point where it coincides with induced DPA investment.

Point *B* on the horizontal axis is the point of balance corresponding to points *A*, *B*, or *C* in Figure 3, i.e., to the points of optimal allocation of the country's resources from the point of view of static equilibrium.[10] If the SOC/DPA ratio is lower than *OB*, we are in the realm of shortage and of induced SOC investment; if it is higher, we are in that of excess capacity and of induced DPA investment. At first sight, it would appear that the greater the departure of the SOC/DPA ratio from *OB*, the greater will be induced investment in either SOC or DPA; and this situation is indeed portrayed up to points *A* and *C* in Figure 4a. The flattening out of the curve showing induced DPA investment is easily explained; obviously this particular inducement effect is subject

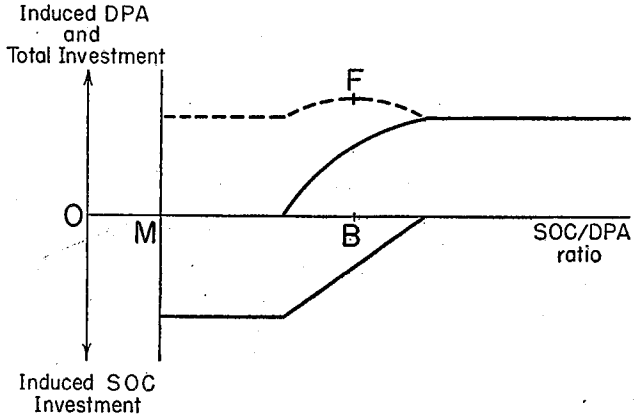

Fig. 4b. Induced investment, with foresight

to decreasing returns as SOC excess capacity rises; within any finite period traffic is not likely to expand in proportion to the number of lanes of a highway. The situation is different for the way in which an SOC shortage induces additional SOC in the next period; here there is no reason for assuming a flattening out. The greater the shortage the greater will be the pressures to remedy it and the greater the amount of capital required to accomplish this end. The limitation is rather a technical one: while the excess capacity of SOC can be as high as we wish to imagine it, there are technological limitations to

10. The value of the ratio does not need to be the same at different levels of output.

a reduction of SOC if a given DPA output is to be maintained. This minimum is shown on our figure as point *M*.

Another limitation comes into play if we think—as we must—in terms of a finite reaction period: then there is a limit to the amount of investment that can be started during that period, regardless of the strength of the inducement. This consideration applies of course also to the induced DPA investment and it results in the two segments of the inducement curve *A'ABCC'* in Figure 4a that are parallel to the horizontal axis.

We may now engage in some speculation about the strength of the inducement effects, as reflected in the shape of the inducement curves. First, let us notice that the curve for *total* induced investment as drawn in Figure 4a has its minimum at the point of balance, *B*. In other words, when the SOC/DPA ratio is such that there is neither excess capacity nor shortage of SOC, and when SOC and DPA are thus in balance, and resources have been optimally allocated between these two categories, then we would actually obtain zero induced investment in the next period. Thus, the contention that balanced growth may be undesirable, besides being unattainable, is illustrated and may well hold true for many underdeveloped countries.

Naturally we can vary our chart to come to quite different conclusions. It is fully conceivable that, once an economy has been progressing over an extended period of time, there will be an overlapping of induced DPA and SOC investments over a significant stretch, a situation shown in Figure 4b. DPA investors who can look back on a long experience of SOC investment being taken care of will be induced to invest prior to the actual point when excess capacity appears; and SOC investment will similarly be undertaken prior to the appearance of an actual shortage. Thus, when the community has acquired the foresight and confidence in further growth characteristic of a progressive economy, the point of balanced growth may also be the point of maximum inducement of further investment, as is shown by point *F* on the *total* inducement curve in Figure 4b. Under these conditions, balanced growth would therefore be desirable not only from the static but also from the dynamic point of view; moreover, it would no longer be unattainable as our strictures against balanced growth on this count would clearly not apply to the advanced type of economy we are now talking about. Regardless of the existence of a

true maximum at *B*, a situation where there is substantial overlapping of the two inducement effects sharply reduces the attractiveness of either variety of unbalanced growth from the point of view of inducing further investment.

We are getting close here to the traditional notions of dynamic equilibrium and induced investment: even though, with SOC and DPA in balance, there exist for the moment neither tensions, pressures, nor special incentives, further balanced growth of these two components of total investment is induced by the anticipation of such pressures and incentives; this anticipation must, in turn, be traced to past experience, which has taught SOC and DPA operators to expect trouble or opportunity, respectively, before either actually arises. In this sense, *it is the experience of unbalanced growth in the past that produces, at an advanced stage of economic development, the possibility of balanced growth.*

But let us return from this El Dorado to our underdeveloped areas, where balanced growth is both unattainable and undesirable. This brings us back to Figure 4a and to speculation about the comparative strength of inducement via SOC shortage or SOC excess capacity.

A basic difference between these two sequences is the type of inducement that is set up. Excess SOC capacity is essentially permissive; while it certainly serves to reinforce motivations that already exist, and may therefore mean the difference between a large flow of DPA investment and a trickle, it invites rather than compels. The opposite holds with respect to the inducement via shortage. A shortage that is experienced as such is bound to lead to attempts to remedy it on the part of those who suffer from it or who stand to gain from its elimination. In situations where motivations are deficient, it therefore seems safer to rely on development via shortage than on development via excess capacity. In other words, if we endow an underdeveloped country with a first-class highway network, with extensive hydroelectric and perhaps irrigation facilities, can we be certain that industrial and agricultural activity will expand in the wake of these improvements? Would it not be less risky and more economical first to make sure of such activity, even though it may have to be subsidized in view of the absence of adequate transportation and power, and then let the ensuing pressures determine the appropriate outlays for SOC and its location? As examples of this type of sequence, one may cite the de-

velopment of Japan, Turkey and, to a considerable extent, of the U.S.S.R.

As already pointed out, the limits to such a policy are set by technological factors, since a minimum amount of SOC is a prerequisite at any given level of DPA. These limits may sometimes be rather narrow, especially in largely undeveloped regions and areas of new settlement where little or no social overhead capital has as yet been created. Naturally, at this stage, only development via excess capacity is feasible; the points *M* and *B* of Figure 4a are merged at the point of origin. The idea that expansion of SOC must under all circumstances precede an expansion of DPA probably originates in the identification of all development situations with such *initial* growth sequences and in the consideration of SOC as a kind of fixed-coefficient "input" that would always have to come into existence prior to "its" DPA-output. But this is surely not a realistic picture for most of today's underdeveloped countries.

The argument in favor of development via SOC shortage applies with particular force to an underdeveloped country's own backward areas. Here we have regions that have, so to speak, hardened in their reluctance to develop and it seems rather unlikely that a purely permissive system—which is moreover a very expensive one—will be effective. In other words, the *BC* segment of the inducement curve—the one corresponding to excess capacity in SOC—is likely to be pitifully flat.

To place one's faith in purely permissive sequences and to rely on the ability of SOC facilities to call forth other economic activities can, under these circumstances, be just as irrational as the so-called "Cargo Cult" that has been engaged in by some of the New Guinea tribes after the lamented departure of the Allied expeditionary force at the end of World War II: "Those in the coastal villages have built wharves out into the sea, ready for the ships to tie up, and those in land villages have constructed airstrips out of the jungle for the planes to land. And they have waited in expectancy for the Second Coming of the Cargoes." [11] Touching as it is, such a belief in the propitiatory powers of social overhead capital should not be the basis of development policy.

11. Raymond Firth, *Elements of Social Organization* (New York, Philosophical Library, 1951), p. 112.

On the other hand, many underdeveloped but developing countries also have their boom towns and their privileged provinces. Here, of course, a policy of providing for plentiful SOC will lead to beneficial repercussions since in such regions many entrepreneurs are always on the brink of investment decisions and will therefore be pushed over it by an improvement in cost and revenue outlook such as is provided by good SOC facilities. On the other hand, a moderate SOC shortage is not likely to do too much damage to a really dynamic developing area. In such a situation industries will think nothing of bringing in their own diesel generators, of digging for their own water, and of building their own access roads and workers' houses. Thus the *BC* curve is not only likely to surge nicely upward, but may even overlap the *AB* curve as in Figure 4b. In other words, a region that is really booming will be much helped by a policy concentrating on SOC, but will not be decisively harmed by its absence.

Observation certainly bears this out. Rapidly growing centers will usually suffer shortages, sometimes because of lack of proper planning, but often also because it would be illegitimate and wasteful to expand SOC facilities in anticipation of the kind of extremely rapid economic progress that does hit a city or area sometimes, but whose occurrence or continuation can never be predicted with confidence. When these shortages do occur, they do not seem to affect the growth perceptibly, but rather are taken as an additional proof that dynamic development is indeed under way. In an underdeveloped country it is often the city with the worst water, power, and housing shortages that is most favored by private investors.

The foregoing is more valid for domestic than for foreign investment. The latter can choose among dynamic centers of several countries and is likely to give some preference to those where shortages are least harassing. This consideration also applies within a country that has several developing centers competing with one another. In such a situation the positive response to excess SOC may be high for one center because of the competition of the others.[12] In other words, the choice of DPA investors is likely to be made primarily on the basis of

12. The situation is a close analogue to that of price elasticity of demand for the export article of a country when this country is one among many producers. That elasticity, as is well known, is always high even though the elasticity for the commodity as a whole may be low.

the comparative SOC endowments of the different "candidate" areas: [13] the *BC* curve is bulging, and development via excess capacity is definitely the choice in this case.

Limits to development via SOC shortage. Our advocacy of development via SOC shortage in certain situations has been hedged by a number of qualifications. Even so, it will come as a shock to those who feel that the inadequacy of transportation and power has been a principal bottleneck in the growth of many important underdeveloped countries. While some observers attribute this situation to inadequate planning and mistaken priorities in development programs, others account for it more simply and often more convincingly by inflation and its consequences: when a country is in the grip of inflation, economic policy-makers are very prone to attempt to hold prices down wherever the administrative difficulties are not overwhelming. Now the prices of public utility services are among the easiest to hold, and are usually subject to direct control of the authorities. Therefore rates of public utilities such as water, electric power, railroads, streetcars, telephones, etc. usually lag far behind the rise in the general price level, with the double result that wasteful consumption of these services is powerfully encouraged while financing of necessary expansions is rendered difficult if not impossible.

There is no doubt that neglect of public utilities can become a most serious drag on economic progress. We have recognized that there is a minimum SOC/DPA output ratio, i.e., a point where DPA output can only be increased if SOC is expanded. But this does not mean that at this point no further DPA investment can take place at all; rather, total productivity of *aggregate* DPA investment will show no further increases. An individual factory can still be established, and while its output will of course be positive, its *social* marginal productivity will be zero, for the productivity of existing firms is lowered once the new firm claims a share of the already short SOC facilities. This situation may show up in marginal firms going out of business, or in a generally poorer quantity and quality performance of all firms, or in both. In other words, the disequilibrium between SOC and DPA has become so pronounced that it is automatically kept from increasing through a refusal of DPA to expand or through the destruction of existing DPA investment as new investment is undertaken.

13. This is shown by the advertisements of the various state development commissions or boards that usually appear in the *New York Times Annual Economic Review* early in January of each year: these organizations point primarily to power and transportation availabilities, factory space, and more intangible factors such as good labor relations or trained labor force.

Unfortunately it is not easy to recognize the point where development of DPA ahead of SOC becomes self-defeating in this fashion, precisely because no absolute obstacle ever arises to debar additional DPA investment. It is quite possible that some countries with protracted inflations have come close to this point especially when the reluctance to raise rates has been further strengthened by the desire to "squeeze" foreign-controlled companies.

Nevertheless, we cannot condemn altogether one of our two approaches—development via SOC shortages—solely on the ground that, if overdone, it *may* lead to stagnation rather than to growth. For it will do so only in a community whose behavior has become thoroughly irrational and where creative responses have been choked off. But which development approach can be expected to work in such a community? When the authorities responsible for public utilities and other SOC facilities fail to respond to dangers so clear and present and to needs as obvious as those that are signalled by collapsing bridges, derailed trains, and constant power failures, what hope is there that these same authorities would know how to plan the construction of bridges, railroads, and power stations *ahead of demand*, without the grossest waste of resources?

We are trying to formulate here a set of economic development policies on the assumption that underdeveloped countries have to contend with some special difficulties and scarcities; but our task would be impossible if we had to assume complete lack of response to development stimuli. Such behavior is best explained as a temporary victory of those forces which are opposed to paying the price of development—from their point of view such behavior is of course entirely rational.

CHAPTER 6

Interdependence and Industrialization

Backward and Forward Linkage Defined

OUR SEARCH for efficient sequences and for mechanisms that tend to maximize "induced" investment decisions must now move on to the directly productive activities themselves. Here we meet with such well-known development alternatives as agriculture versus industry, export promotion versus import substitution, heavy versus light industry, etc. Rather than examine in isolation each one of these specific problems we shall attempt to sketch a generalized method of attack.

In our discussion of SOC versus DPA we made two assumptions:

1. that within certain limits the proportions in which investment must be divided between these two categories are variable (this assumption resulted in the possibility of excess or shortage of social overhead capital); and

2. that in case of a shortage, the remedy of imports was not available (this is fairly realistic in the case of transportation and power, not to speak of more basic SOC activities such as education and law enforcement).

Both assumptions will now be reversed; insofar as relations among various directly productive activities are concerned, we shall operate on the assumption, familiar from input-output analysis, that input coefficients are fixed or at least that needed inputs increase monotonically with outputs. In other words, while it was conceivable that an economy is inadequately or overgenerously equipped with technical training facilities, port installations, roads, and electric power stations relative to its endowment of manufacturing capacity, it is plausible to suppose (at least with given techniques) that an increase in the output of cotton yardage requires a larger input of raw cotton. Flexibility is restored, on the other hand, by the assumption that if a required input is not supplied from domestic sources it can be imported.

The pattern of pressures, incentives, and repercussions is therefore likely to be quite different from the one that was characteristic of the SOC-DPA relationship.

In discussing the latter, the concepts of SOC shortage and excess capacity emerged as the principal agents of further development. With our new assumptions, a specific shortage of goods and services required as inputs for a certain newly established economic activity is inconceivable. If the required inputs are not available, the activity will simply not take place. If it does take place, the inputs must already be forthcoming—either as a result of domestic production or through imports. Thus it seems that there is little room here for the kind of direct and strong pressure that leads from one productive activity toward another. But important stimuli result nevertheless from the fact that the setting up of an industry brings with it the *availability of a new, expanding market for its inputs* whether or not these inputs are supplied initially from abroad.

Similarly with what we called excess capacity. At first it would seem that this concept has little place within the DPA sector. An industry must sell its output. It will come into being only if it anticipates being able to do so; if the anticipation turns out to be a miscalculation, the industry will have to close its doors. Thus the coming into existence of an industry must be the result of some pre-existing demand, but can it be held to induce new activities and demand? Such repercussions are less obvious than the road which permanently encourages new traffic flows. Nevertheless, the domestic availability of a product whose output will ordinarily be expanded if necessary does have some features in common with the availability of a basic service.

It may seem that, given our assumption that DPA products can be imported, there is no reason why domestic availability of a new product should prove a better spur to further economic activity involving the product as an input than availability from foreign sources of supply that existed all along. In pure theory, this may indeed be so, but in practice three important considerations make domestic availability a considerably more effective spur to further development:

1. Importing requires special skills and therefore reduces the number of potential entrants.

2. Importing is subject to special balance-of-payments uncertainties, and production largely based on imports is therefore particularly

risky; if inflation is expected to proceed more rapidly at home than abroad and adjustment of the exchange rate is held back, quantitative import restrictions become likely; if, on the other hand, the exchange rate is allowed to depreciate freely, it is likely to do so faster than the domestic purchasing power of the country's currency and there would be a long-run cost advantage in relying as much as possible on domestic inputs.

3. Finally and perhaps most important, the fact that a certain product is produced domestically is likely to result in efforts on the part of the producers to propagate its further uses and in their financial participation in such ventures. The domestic availability of a product thus brings into being active forces that make for its utilization as input in new economic activities catering to newly "entrained wants." In this respect, it is therefore a less purely permissive inducement mechanism than the existence of a road which merely "invites" more traffic.

Thus, in close analogy to the alternative between development via shortage and development via excess capacity which we described for the SOC-DPA situation, two inducement mechanisms may be considered to be at work within the DPA sector:

1. The input-provision, derived demand, or *backward linkage effects,* i.e., every nonprimary economic activity, will induce attempts to supply through domestic production the inputs needed in that activity.

2. The output-utilization or *forward linkage effects,* i.e., every activity that does not by its nature cater exclusively to final demands, will induce attempts to utilize its outputs as inputs in some new activities.

Development policy must attempt to enlist these well-known backward and forward effects; but it can do so only if there is some knowledge as to how different economic activities "score" with respect to these effects. Ordinarily economists have been content with general references to the advantages of external economies, complementarities, cumulative causation, etc. But no systematic effort has been made to describe how the development path ought to be modified so as to maximize these advantages even though the existence of input-output statistics supplies us with a few tools for an analysis of this kind.

First, a further note on the linkage concept itself. What do we imply when we speak of the linkage effects emanating from industry *A* toward industry *B*? Language can be quite ambiguous here, for we may have in mind the potential *importance* of the linkage effect in terms of, say, the net output of the new industries that might be called

forth; or we may mean the *strength* of the effect, i.e., the probability that these industries will actually come into being. The total effect could be measured by the sum of the products of these two elements; in other words, if the establishment of industry W may lead, through linkage effects, to the establishment of n additional industries with net outputs equal to x_i $(i = 1,2...n)$ and if the probability that each one of these industries will actually be set up as a result of the establishment of industry W is p_i $(i = 1,2...n)$, then the total linkage effect of industry W is equal to $\sum_{1}^{n} x_i p_i$.

The probabilities can be interpreted as measuring the strength of the stimulus that is set up. For backward linkage, this strength can be roughly measured as follows: suppose industry W requires annual inputs of $y_1, y_2...y_n$ and suppose that the minimum economic size (in terms of annual productive capacity) of firms that would turn out these inputs is equal to $a_1, a_2...a_n$; then the strength of the stimulus or the probability that the setting up of industry W will lead to the setting up of industries producing the inputs is equal to the ratio of the y's to the a's.[1] Minimum economic size is not a technical concept, but is defined in economic terms relative to normal profits and efficient foreign suppliers. In other words, it is the size at which the domestic firm will be able both to secure normal profits and to compete with existing foreign suppliers, taking into account locational advantages *and* disadvantages as well as, perhaps, some infant industry protection. In this way comparative cost conditions are automatically taken into account.[2]

In the case of forward linkage, an interpretation of the p's is less

1. The ratio is to be defined as having a ceiling of 1, i.e., the value of the ratio is equal to unity, whenever the y's are equal to *or larger than* the a's. Note also that the y's are equivalent to the gross output of the new industries or firms in physical terms whereas the x's are their net outputs in value terms.

2. Data on the economic size of plants in different industries would be the starting point for determining minimum economic size in different countries. Research in this area in relation to economic development is surprisingly scant, except for the pioneering article of K. A. Bohr, "Investment Criteria for Manufacturing Industries in Underdeveloped Countries," *Review of Economics and Statistics,* 36 (May 1954), 157–66. Some basic data for small and medium plants are available in a series of Industrial Reports issued by the Office of Industrial Resources of the International Cooperation Administration. For data on optimum plant size in U. S. industry, see J. S. Bain, *Barriers to New Competition* (Cambridge, Mass., 1956), ch. 3.

straightforward. The concept of economic size is not helpful here, since the size of the market for the industries that might be brought into being through forward linkage does not depend on their suppliers. A clue can perhaps be found in the importance of the articles produced by industry W as inputs for the output of the to-be-linked industry. If these inputs are a very small fraction of the industry's eventual output, then their domestic availability is not likely to be an important factor in calling forth that industry. If, on the other hand, these articles are subjected to few further manufacturing operations, then the strength of the forward stimulus is likely to be substantial, provided demand is sufficient to justify domestic production.

In these cases, then, importance and strength—the x's and the p's—of the linkage effect are inversely correlated. Industries where the x's are small and the p's large are sometimes aptly called "satellite" industries. They are almost unfailingly established in the wake of industry W but are of minor importance in comparison to that industry. Thus defined, satellite industries can be established through backward or forward linkage. In the case of cement, for instance, the manufacture of multi-wall bags for packing purposes represents backward linkage while the establishment of a cement block industry represents satellite formation through forward linkage. A satellite industry usually has the following characteristics:

a. it enjoys a strong locational advantage from proximity to the master industry;

b. it uses as principal input an output or by-product of the master industry without subjecting it to elaborate transformation, or its principal output is a—usually minor—input of the master industry; and

c. its minimum economic size is smaller than that of the master industry.

While satellite industries are almost certain to be established once the master industry is in place, the establishment of industry W also results in stimuli toward the setting up of nonsatellite industries. In these cases, the strength of the stimulus is infinitely weaker, but the stake is far bigger. Examples of such a situation are the stimulus that the setting up of a multi-wall bag factory gives toward the creation of a pulp and paper industry or, for the case of forward linkage, the stimulus given by the establishment of an iron and steel industry to

all the metal-fabricating industries. Here the establishment of one industry is a contributing factor which by itself is quite unlikely to result in the creation of the others; but when we speak of external economies and complementarities, we think at least as much of these uncertain linkages as of the far more certain, but also far less significant, satellites with which any industry of a certain size surrounds itself. The weakness of the stimulus in the case of nonsatellites can be explained by the absence of the three factors that define satellites. Linkage is reduced to the fact that an input of the newly established industry is an output of the to-be-created industry or vice versa, but the established industry would not be the principal customer or supplier of the to-be-created industry; in fact, particularly in cases of backward linkage, minimum economic size of the to-be-created industry would frequently be larger than that of the industry where the linkage originates.[3]

In spite of the importance of the nonsatellite type of linkage, it seems necessary to provide for some arbitrary cut-off point for small probabilities. It is all very well to say that the establishment of a brewery sends out a stimulus in the direction of a paper industry because of the labels needed for the beer bottles, but by itself this stimulus is not likely ever to lead to the setting up of a paper mill. Thus, if we consider *in isolation* the linkage effects exclusively of the beer industry on further industrial or agricultural development, we should consider only those stimuli whose probability exceeds a certain critical value, say one-half.[4]

If we proceed in this way, the joint linkage effects of two industries, say beer and cement, considered as a unit, are likely to be larger than the sum of their individual linkage effects, since some of the xp products which are omitted in computing the individual effects because

3. To the extent that the minimum economic size of an industry is larger the farther away one moves from the finished consumer or producer goods stage. This is, of course, by no means universally true as is shown, e.g., by the large minimum size of the motor vehicles makers.

4. It is a good rule of thumb that an industry can properly be established in an underdeveloped but developing country as soon as existing demand is equal to one-half of the economic size of the plant as defined above. The additional demand needed to justify the investment can be expected to come from the growth of existing demand and from the development of new demand through forward linkage, once the plant is in existence.

the p's are below the critical value will exceed this value if added together for both industries. Here we have an argument in favor of multiple development that we would consider convincing were it not that our principal argument against it is concerned with its feasibility rather than with its desirability.

The fact that the linkage effects of two industries viewed in combination are larger than the sum of the linkage effects of each industry in isolation helps to account for the cumulative character of development. When industry *A* is first set up, its satellites will soon follow; but when industry *B* is subsequently established, this may help to bring into existence not only its own satellites but some firms which neither *A* nor *B* in isolation could have called forth. And with *C* coming into play some firms will follow that require the combined stimuli not only of *B* and *C* but of *A*, *B*, and *C*. This mechanism may go far toward explaining the *acceleration* of industrial growth which is so conspicuous during the first stages of a country's development.

A Mental Experiment

Let us attempt to take a closer look at the cumulative effect of industrialization. The problem is now simply formulated: how are linkage effects maximized? Obviously, even if we had the necessary quantitative information about these effects, an answer to this question would yield only *one* investment criterion among several others. Therefore, the usefulness of the criterion lies in the guidance it could provide in specific situations with respect to alternatives that arise in the course of development, rather than in the design of any rigid sequence that might be derived from it.

As a first step, it is instructive to look at the various sectors of an economy with the aim of appraising the amount and kinds of linkage effects which they exert. With linkage most fully developed in advanced industrial countries, it is tempting to turn to them to discover those sectors or subsectors that ought to receive high ratings in any development program, not only on account of the intrinsic usefulness of their output, but because of the further development stimuli which are likely to emanate from them.

What use can be drawn from the statistical results of input-output analysis in this connection? We may, of course, measure the degree of interdependence shown by any one industry by computing

1. the proportion of its total output that does not go to final demand but rather to other industries, and
2. the proportion of its output that represents purchases from other industries.

But these two measures of the extent to which any one industry interlocks with others within a national economy can be taken to represent forward and backward linkage effects only on the basis of a mental experiment: we would have to imagine for every industry in turn that the country's development started with it, so that all its sales to and purchases from other domestic industries are fancied to have developed as a sequel to its foundation. In fact, of course, an industry *W* which shows the highest degree of interdependence could very well have been set up last, thus proving that maximum interdependence is quite compatible with complete absence of active linkage effects. Nevertheless, *if W* had been set up first, and assuming that the total output and its commodity composition would have come out to the same final result shown in the input-output table, then the industries that are using *W*'s products in their own manufacturing processes, and those that are suppliers of *W*, would have received important growth stimuli from it. Since we are interested in appraising the comparative importance of such hypothetical stimuli in different industries to help us find, precisely in a nonindustrial country, the most efficient sequence for its industrialization, the results of this mental experiment are pertinent to our inquiry.[5]

In a recent study by Chenery and Watanabe,[6] the degree of interdependence of various industries has been computed and averaged for Italy, Japan, and the United States, and the resulting figures reproduced in the table below can be taken as a general indication of the ranking of these industries from the point of view of backward and forward linkage effects. Because of the proviso contained in note 5, averages for countries with different industrial structures are perhaps better suited to provide such indications than individual country figures.

5. On the condition that we expect the commodity composition of the underdeveloped country's output to bear eventually some resemblance to that of the country on whose input-output statistics we perform the experiment.

6. H. B. Chenery and T. Watanabe, "International Comparisons of the Structure of Production," paper presented at the Cleveland Meeting of the Econometric Society, Dec. 1956 (dittoed), to be published in *Econometrica.*

Average Degree of Interdependence of Economic Sectors in Italy, Japan, and the United States

	Interdependence through Purchases from Other Sectors [a] *(Backward Linkage)*	*Interdependence through Sales to Other Sectors* [b] *(Forward Linkage)*
1. *"Intermediate Manufacture"* (backward and forward linkage both high)		
Iron and Steel	66	78
Nonferrous Metals	61	81
Paper and Products	57	78
Petroleum Products	65	68
Coal Products	63	67
Chemicals	60	69
Textiles	67	57
Rubber Products	51	48
Printing and Publishing	49	46
2. *"Final Manufacture"* (backward linkage high, forward linkage low)		
Grain Mill Products	89	42
Leather and Products	66	37
Lumber and Wood Products	61	38
Apparel	69	12
Transport Equipment	60	20
Machinery	51	28
Nonmetallic Mineral Products	47	30
Processed Foods	61	15
Shipbuilding	58	14
Miscellaneous Industries	43	20
3. *"Intermediate Primary Production"* (forward linkage high, backward linkage low)		
Metal Mining	21	93
Petroleum and Natural Gas	15	97
Coal Mining	23	87
Agriculture and Forestry	31	72
Electric Power	27	59
Nonmetallic Minerals	17	52

	Interdependence through Purchases from Other Sectors [a] *(Backward Linkage)*	*Interdependence through Sales to Other Sectors* [b] *(Forward Linkage)*
4. *"Final Primary Production"* (backward and forward linkage both low)		
Fishing	24	36
Transport	31	26
Services	19	34
Trade	16	17

a. Ratio of interindustry purchases to total production (%).
b. Ratio of interindustry sales to total demand (%).

Source: Chenery and Watanabe, "International Comparisons," p. 11. Reproduced by permission of the authors.

The four categories in the table have been established by Chenery and Watanabe, and they are meaningful within our own conceptual framework. Within each category, the industries have been ranked according to their combined score (backward plus forward linkage). As we attribute more importance to backward than to forward linkage,[7] we place industries with high backward and low forward linkage ahead of those that have the inverse characteristic.

The ranking does an injustice to machinery and also to transport equipment. The low value of forward linkage in their case is probably due to the fact that sales of these industries to other industries are construed in input-output tables as final demand deliveries because they are accounted for under capital formation. Looking at interindustry transactions from the point of view of linkage effects, there is no essential difference between the stimulus toward setting up an insecticide mixing plant and that toward installing a tractor assembly plant which emanates from insecticide or tractor purchases by agricultural establishments.

In some cases, the mental experiment through which we identify interdependence with linkage effects takes on a somewhat eerie flavor. For instance, the largest value for backward linkage is found for grain mill products, but it is highly unrealistic to think of wheat and rice cultivation as being "induced" by wheat and rice mills; rather the mills must be regarded as satellites (through forward linkage)

7. See p. 116.

of the foodcrop-growing activities. Nevertheless, even in this instance, backward linkage is sometimes experienced in underdeveloped countries as will be seen below.

In any event it is interesting to note that the industry with the highest combined linkage score is iron and steel. Perhaps the underdeveloped countries are not so foolish and so exclusively prestige-motivated in attributing prime importance to this industry!

The interdependence ratios of the table are very rough indexes of the potential linkage effects that might be introduced into nonindustrial economies by specific industrial sectors. A more refined measure of backward linkage can be obtained by considering the inverse of the input-output matrix. This inverse matrix makes it possible to estimate the *direct and indirect* repercussions of an increase in final demand requirements for any one industry on the other sectors of the economy. Since indirect repercussions are not taken into account when one computes simply the ratio of an industry's purchases from other industries to the total value of its output, the measure derived from the inverse matrix is more comprehensive. A Danish economist has proposed this measure—he calls it "power of dispersion"—as one way of identifying "key industries," and has argued that knowledge of these measures could be of value in a depression because it would permit us to focus recovery policy on those industries whose expansion would "lead to a general increase in economic activity embracing all or at least most industries." [8] In addition, we believe that computation of these indexes may be of interest not only to those who seek to *re*activate a developed economy but also to those who attempt to activate an underdeveloped economy.

The knowledge of the approximate ranking of an industry from the point of view of forward and backward linkage effects as derived from existing developed economies through their input-output tables is, I believe, useful to the economist-planner in underdeveloped areas. It is something to be added to his criteria-box. But excessive reliance should obviously not be placed on these rankings, based as they are on a mental experiment subject to numerous qualifications. Industrial development clearly cannot be started everywhere with an iron and

8. P. N. Rasmussen, *Studies in Inter-Sectoral Relations* (Copenhagen, Einar Harcks, 1956), p. 141. Unfortunately the empirical studies included in the book do not do justice to Rasmussen's very interesting analytical tools because of excessive aggregation. Thus, all manufacturing is brought together in a single sector!

steel industry just because this industry maximizes linkage. It is far more useful to look at the structure of underdeveloped countries and to examine how linkage effects normally make their appearance: such an analysis is likely to yield some hints about the possibility of influencing development in such a way as to strengthen these effects.

Backward Linkage at Work

The lack of interdependence and linkage is of course one of the most typical characteristics of underdeveloped economies. If we had homogeneous input-output statistics for all countries, it would certainly be instructive to rank countries according to the proportion of intersectoral transactions to total output; it is likely that this ranking would exhibit a close correlation with both income per capita and with the percentage of the population occupied in manufacturing.

Agriculture in general, and subsistence agriculture in particular, are of course characterized by the scarcity of linkage effects. By definition, all *primary* production should exclude any substantial degree of backward linkage although the introduction of modern methods does bring with it considerable outside purchases of seeds, fertilizers, insecticides, and other current inputs, not to speak of machines and vehicles. We may say that the more primitive the agricultural and mining activities, the more truly primary they are.

Forward linkage effects are also weak in agriculture and mining. A large proportion of agricultural output is destined directly for consumption or export; another important part is subjected to some processing in industries that can be characterized as satellite inasmuch as the value added by them to the agricultural product (milling of wheat, rice, coffee, etc.) is small relative to the value of the product itself. Only a comparatively small fraction of total agricultural output of underdeveloped countries receives elaborate processing, which usually takes place abroad.

The case for inferiority of agriculture to manufacturing has most frequently been argued on grounds of comparative productivity. While this case has been shown not to be entirely convincing,[9] agriculture certainly stands convicted on the count of its lack of direct stimulus

9. Viner, *International Trade and Economic Development* (Glencoe, Ill., Free Press, 1952), pp. 63–73; and Buchanan and Ellis, *Economic Development*, pp. 259–63.

to the setting up of new activities through linkage effects: the superiority of manufacturing in this respect is crushing. This may yet be the most important reason militating against any complete specialization of underdeveloped countries in primary production.

The grudge against what has become known as the "enclave" type of development is due to this ability of primary products from mines, wells, and plantations to slip out of a country without leaving much of a trace in the rest of the economy. Naturally hostility to the profits earned by foreign companies plays an important role in such attitudes; but the absence of direct linkage effects of primary production for export lends these views a plausibility that they do not have in the case of foreign investment in manufacturing. I say plausibility rather than validity, for while as such the primary production activities leading to exports may exert few developmental effects, they do finance imports which can become very powerful agents of development as we shall see below.

Since interdependence in the input-output sense [10] is so largely the result of industrialization, we must now attempt to trace the various ways in which manufacturing and the accompanying linkage effects make their appearance. In this connection, we shall utilize another one of Chenery's findings, namely that more than ninety percent of all input-output flows can usually be arranged in a triangular pattern.[11] Circularity—i.e., the fact that coal is needed for steel-making and steel for coal mining—is undoubtedly present in the structure of a country's production, but apparently to a much smaller degree than would be suspected upon looking at an input-output table that has not been "triangularized." In other words, there is no compelling *technological* requirement for the simultaneous setting up of various industries, an interesting complement to our case against the existence of such a requirement on economic grounds.

In a triangular arrangement of the input-output matrix, there is a "last" sector whose output goes entirely to final demand and which takes in inputs from a number of other sectors; the second-to-last sector sells its output to final demand and to the last sector and buys

10. This qualifying clause is needed, as it is possible to have extensive division of labor, and therefore interdependence in final demand, in an economy that shows very little "structural" interdependence.

11. Chenery and Watanabe, "International Comparisons," pp. 7–10.

inputs from some or all other sectors except from the "last"; and so on, until we come to the "first" sector whose output goes to all the subsequent sectors and possibly also to final demand, but which does not use any inputs from other sectors.

Industrialization can of course *start* only with industries that deliver to final demand, since *ex hypothesi* no market exists as yet for intermediate goods. This means that it will be possible to set up only two kinds of industries:

1. those that transform domestic or imported primary products into goods needed by final demands;
2. those that transform imported semimanufactures into goods needed by final demands.

To the pioneer industrial countries only the first course was open, and this explains the towering importance of a few industries (textiles, iron and steel, pottery) during the early stages of the Industrial Revolution. In today's underdeveloped countries the textiles, food processing, and construction materials industries based on local materials are still of great importance, but, to a very significant extent, industrialization is penetrating these countries in the second manner, through plants that perform the "final touches" [12] on almost-finished *industrial* products imported from abroad. Examples are the many converting, assembly, and mixing plants, the pharmaceutical laboratories, the metal-fabricating industries, and many others. This trend has many advantages: it often provides an investment outlet for small amounts of capital that might not easily become available for ventures which require the pooling of the resources of many investors, and it makes it possible to start industrial undertakings without the heavy risk that comes in underdeveloped countries from having to rely on the output of unreliable domestic producers.

In this way underdeveloped countries often set up "last" industries first—i.e., these are "last" industries considering the input-output flow of the advanced countries: what in these countries are inputs from the other sectors are replaced in underdeveloped countries by imports. Such industries could be termed "enclave import industries," in analogy to the enclave export activities that were previously mentioned. For here again we have an undertaking that at least in its

12. This term was used by Hayek in his discussion of the "stages of production." Cf. *Prices and Production* (London, Routledge, 1931), p. 70.

beginning is antiseptically linkage-free; materials are imported from abroad, some value is added to them through mixing, assembling, packaging, etc., and the finished product is rushed to the final consumers. The enclave nature of these industries is sometimes emphasized by the location of the plant at a point as close as possible to the most convenient port of arrival of the imported materials, and again this type of venture has proven particularly attractive to foreign capital—many of the branch plants owned by foreign corporations specialize in this kind of operation.

But there is a considerable difference between the enclave export and enclave import activities. The former have great trouble in breaking out of the enclave situation. Usually some forward linkage effects can be utilized—ores and cane sugar can be refined before being shipped. But the scope for such operations is strictly limited. With respect to import enclave industries, the situation is radically different: they set up backward linkage effects of practically infinite range and depth.

In fact, much of the recent economic history of some rapidly developing underdeveloped countries can be written in terms of industrialization working its way backward from the "final touches" stage to domestic production of intermediate, and finally to that of basic, industrial materials. In this way, industrialization has even proven to be a powerful stimulus to the development of agriculture. By providing a reliable market, processing industries originally based on imported agricultural materials such as cotton textiles and beer have stimulated in Colombia the domestic production of cotton and barley. In Brazil, likewise, "modern cotton, peanut and cereal processing plants sometimes preceded the expansion in agricultural output. Similarly the very sizable investments in cotton, sisal and cocoa ginning and pressing plants, gave evidence of the way in which industrial and commercial capital has been ready to seek out profitable opportunities to strengthen the country's raw material base." [13] Backward linkage effects are important not only from secondary back to primary production, but also from tertiary back to both secondary and primary production. The experience of Sears, Roebuck and of some

13. *The Development of Brazil,* Report of the Joint Brazil-United States Economic Development Commission, Institute of Inter-American Affairs (Washington, 1954), p. 12.

successful supermarket ventures in Latin America are cases in point.[14]

In most of these cases, imported goods have been gradually replaced by domestic production which has been called forth by the existence of a large and stable market. Of considerable importance are the backward linkage effects that are the combined result of the existence of several "last stage" industries. The minimum economic size of many intermediate and basic industries is such that in small markets a variety of user industries needs to be established before their combined demand justifies a substitution of imports of intermediate and basic goods by domestic production.

A Model of Capital Formation Based on Backward Linkage

In view of the importance of backward linkage as a development mechanism, it may be interesting to construct a growth model where this mechanism would act as a prime mover.

As backward linkage is brought into play by increases in demand, we shall suppose that autonomous growth is taking place as a result of some net capital formation, improvement in efficiency, and rise in exports. But let this be a slow growth, far below the ceiling set for an economy by its potential ability to generate savings and to attract foreign capital. We then assume that backward linkage unfailingly triggers additional net capital formation whenever the imports of some commodity pass the "threshold" of minimum economic size. The situation may be analyzed in the following terms using familiar input-output concepts:

Let there be n activities the first k of which are not carried on within our country at the beginning of the development process. The outputs of these activities which are carried on abroad are imported into the country to be used (1) as inputs in the n–k activities that are carried on within the country; and/or (2) directly as final demand deliveries. This yields an input-output table which differs from the usual type in that imports have been disaggregated. If we assume that

14. R. Wood and V. Keyser, *Sears, Roebuck de México,* S.A., National Planning Association, Washington, 1953. Experience of Sears, Roebuck in other Latin American countries has been similar to that described in this pamphlet. The International Basic Economy Corporation (IBEC) of New York has reported favorable effects of its supermarkets on agricultural production in Venezuela.

imports take place only in commodities that are not produced domestically, then we can represent the table as follows:

		Intermediate Demands from: 1 .... k	$k+1$ ..... n	Final Demand	Total Demand
Outputs from	1	0 0	$M_{1,k+1} \cdots M_{1n}$	M_{1F}	M_1
	2	0 0	$M_{2,k+1} \cdots M_{2n}$	M_{2F}	M_2
	k	0 0	$M_{k,k+1} \cdots M_{kn}$	M_{kF}	M_k
	$k+1$	0 0	$X_{k+1,k+1} \cdots X_{k+1,n}$	$X_{k+1,F}$	X_{k+1}
	n	0 0	$X_{n,k+1} \cdots X_{nn}$	X_{nF}	X_n
Value added		0 0	$X_{V,k+1} \cdots X_{Vn}$	—	—
Total input		0 0	$X_{k+1} \cdots X_n$	—	$X+M$

The first k columns of the matrix are filled with zeros because they correspond to the inputs of the first k activities which, in accordance with our assumption, have not as yet come into being within the country considered.

We are interested in the growth of the various import categories through time and in the consequences of their growth for capital formation. The imports M_1, $M_2 \ldots M_k$ are determined by specifying final demands; this gives us directly the final demand component of the M's (M_{1F}, $M_{2F} \ldots M_{kF}$), and indirectly, via the inverse matrix, the intermediate import demand components.

Suppose that we know in this way the growth of the M's through time, and that we know also for each M the domestic production threshold, i.e., the minimum economic size at which domestic production is undertaken. Let the outputs defining these thresholds be T_1, $T_2 \ldots T_k$, and the corresponding capital requirements K_1, $K_2 \ldots K_k$. At the beginning of the growth process all the M's are smaller than the corresponding T's. But with the growth in final demands, a point will come at which some M_j will be equal to or larger than T_j and thereupon the economy will make the investment K_j. In each period, induced capital formation is therefore equal to the sum of those K's for which this occurs. In this way, one can derive an induced investment pattern that is likely to exhibit strong ups and downs even on the assumption of a perfectly smooth path for the growth of demand.

It is conceivable that this model could be of considerable help in accounting for the sudden spurt of investment relative to income—Rostow's "take-off" point—that appears to characterize the growth process of a number of countries. On our assumptions, a sudden spurt could easily occur in newly industrializing countries. But once domestic production has been established in a large number of lines, capacity can be adjusted far more gradually to increasing output. Therefore, as the industrial base of a country expands, capital growth in response to gradual increases in final demands is likely to be steadier than in the early stages of a country's development.

Empirical studies designed to determine the probable capital requirements of an economy that would develop in accordance with the model just outlined could be of considerable interest. Such studies might permit us to discover certain typical oscillations of the investment impulses which underdeveloped countries are likely to receive along their growth path.

The model could conceivably be put to a different and more intriguing use. Instead of estimating how much capital formation is "triggered" when the final demands are rigidly specified in advance, we could think of influencing capital formation by modifying observable trends in the growth of final demands. For instance, we could wish to maximize backward linkage effects and the concomitant capital formation, and we would then want to manipulate final demands with this end in view. Such manipulation would, of course, have to remain subject to certain reasonable constraints such as:

1. that there is a ceiling for the annual rate of growth of total imports; in view of the assumption of fixed coefficients, this condition automatically places a ceiling on the growth rate of all domestic outputs into which imports enter at all;

2. that the departure of the *manipulated* final demands in either direction from the *expected* demands (i.e., those that would obtain without interference) are kept within certain boundaries.

Whether or not such a maximization problem can receive a general solution, some interference, through tariffs, excise taxes, and subsidies, with the developing consumption of a country may be justified if it can be demonstrated that a certain growth pattern of consumption would exert far more powerful backward linkage effects than the pattern that is likely to develop in the absence of interference. This holds even though one discards the unrealistic assumption that invest-

ments will necessarily and always be undertaken as soon as the threshold is passed. The important point remains that investment decisions are made much easier once this is the case.[15]

To give one example: the minimum size of an automobile body stamping plant varies sharply for different types of automobile. It is much higher for today's typical American car with its huge fenders that merge with the body than for automotive vehicles requiring smaller and simpler stampings. By fiscal policies that favor the latter type of vehicle a country may be able to advance by several years the point at which the establishment of a stamping plant within its territory becomes feasible and attractive to investors.

The rationale for interference with the market mechanism and consumers' preferences which we have just given is particularly strong in slow-moving economies where industrial growth is incipient. If, in such countries, capital formation can be called forth merely by rearranging and concentrating the pattern of imports, then such interference may often be deemed a price well worth paying for an increase in the country's pace of industrialization. In rapidly developing countries, on the other hand, where inflation is rampant, we may use our knowledge of these backward linkage effects for the opposite purpose: namely to ward off the bunching of investment demands that might occur at certain stages of the development process as a result of too many "thresholds" being crossed at the same time.

Combining Backward and Forward Linkage

Backward linkage effects are much neater than forward linkage effects. If it is rather daring to assume that a certain investment will take place as soon as domestic demand reaches the "threshold," it would be downright absurd to set up any model that would presume to indicate which kind of metal-fabricating industries would come into existence at what point of time in the wake of the establishment of a basic iron and steel industry. As was already indicated earlier in this chapter, forward linkage could never occur in pure form. It must always be accompanied by backward linkage, which is the result of the "pressure of demand." In other words, the existence or anticipation

15. Kafka has called attention to the related idea that the distribution of income will affect the pattern and timing of a country's industrialization; see "Some Aspects of the Theoretical Interpretation of Latin-American Economic Development," paper given at the Rio Roundtable of the International Economic Association, 1957, p. 19 (mimeographed). For any one level of income there will always be commodities whose domestic production could be started, and others whose production would cease to be profitable, if the distribution of income were more equal, or more unequal, than it actually is. Theoretically, it is therefore possible to define, at any given level of income, a distribution of that income which maximizes backward linkage effects.

of demand is a condition for forward linkage effects to manifest themselves.

While forward linkage cannot therefore be regarded as an independent inducement mechanism, it acts as an important and powerful reinforcement to backward linkage for the reasons listed earlier in this chapter. Investment decisions that are taken as a result of both backward and forward linkage are caught, as it were, in a pincer movement and must be prized by us since they are sure to be particularly easy-to-take ones.[16]

How are such pincer movements engineered in the course of economic development? They are somewhat difficult to visualize on the basis of the traditional concept of "stages of production" where the successive stages are farther and farther removed from final consumption. But this concept is unrealistic as has been shown by input-output analysis. Many industries produce intermediate goods for other industries and serve final demand at the same time. Thus it is quite possible for industry *A* to be established as a result of final demand for its products crossing the threshold, and then for *B* to follow suit not only because of demand factors but also because *B* intends to use *A*'s products as a principal input. Such a development has a particularly dynamic quality because it necessitates an expansion of industry *A*, which was originally set up only in response to final demand and must now satisfy new industrial customers as well. In other words, while the existence of industry *A* helps to induce the establishment of industry *B*, this establishment in turn induces the building of new capacity for *A*.

This kind of *pincer cum feedback* effect can only be obtained with the help of industries that, in the triangularized matrix of interin-

16. The approach toward the establishment of an industry through backward and forward linkage illustrates the similarity of economic development to a jigsaw puzzle—if a piece is surrounded from several sides, it is easier to fit it in. See p. 82.

dustry transactions, are located at some distance from the top rows. This means intermediate or "basic" industries whose products are distributed as inputs through many other industrial sectors besides also going directly to final demand. It is clear now that such industries should be given preference over the "last" industries, if they are at all economically feasible.

There are other reasons why the most "efficient" or "dynamic" way to work one's way through the triangularized matrix may not be by a gradual trickling down from the top. The ability of underdeveloped countries to start industrialization in this fashion by giving the "last touches" to imported materials is no doubt an advantage inasmuch as it permits industries to be started even in areas where markets are small and technical knowledge and organizational know-how are scarce. But it is also a disadvantage, for it builds up resistances for every new step in the trickling-down process. In dealing with backward linkage effects, we have thus far taken it almost for granted that as soon as domestic demand passes the threshold of minimum economic size, domestic production will be undertaken. But while some forces no doubt make for this course, counterforces are also at work. The industrialist who has worked hitherto with imported materials will often be hostile to the establishment of domestic industries producing these materials. First, he fears, often with good reason, that the domestic product will not be of as good and uniform quality as the imported one. Secondly, he feels that he might become dependent on a single domestic supplier when he could previously shop around the world. Thirdly, he is concerned about domestic competition becoming more active once the basic ingredients are produced within the country. Finally, his location may be wrong once the source of supply of the materials he uses is thoroughly altered.

For all these reasons, the interests of the converting, finishing, and mixing industries are often opposed to the establishment of domestic sources of supply for the products that they convert, finish, or mix. It takes a fairly violent shock—usually resulting from balance-of-payments or inflationary disturbances—to divert such industries from their defense of the status quo.

Therefore, excessive gradualism in introducing industry by successive small bits of value added may not pay off. Whereas the first steps are easy to take by themselves, they can make it difficult to take the

next ones. An industrialization that proceeds in this way is a particularly quiet and uninspired affair, almost at the antipodes of Schumpeter's creative entrepreneur. The founding of a new industry in an advanced industrial country is always a job full of excitement and travail that brings with it new development stimuli in many directions: sources of supply for needed materials must be located or their production in accordance with new specifications must be organized; there is much experimenting with alternative techniques and layouts. When the same industry is set up twenty or thirty years later in an underdeveloped country, the operation in itself is already bound to be far less "exciting," since the technological problems are by then largely solved and the industry is past its phase of fastest technological progress. If, in addition, the industry is of the "enclave import" type, i.e., entirely based on converting imported materials, the national economy is deprived, at least at this stage, of the *unsettling* effects of industrialization that are so beneficial for further development. Borrowing a term from the theory of the multiplier, we may say that these imports of semifinished materials that are always ready to rush in from abroad whenever an industrial project is being considered are real *leakages* of development effects.

Certainly, but for the process of starting with the "last touches," many industries could never be undertaken at all in underdeveloped countries; once this is recognized, however, much is to be said for biting off *as large pieces of value added at a time* as the underdeveloped country can possibly digest.