THEORIES OF ECONOMIC UNDERDEVELOPMENT:
A GENERAL EQUILIBRIUM ANALYSIS

Dissertation
submitted in partial fulfilment of the requirements for
the Degree of DOCTOR OF PHILOSOPHY of Rhodes University

by

PHILIP ANDREW BLACK

DECEMBER 1981
ACKNOWLEDGEMENTS

This thesis is partly the outcome of a long-standing association with my supervisor and colleague, Prof. Mike Truu. Over the years I have benefited greatly from his extensive knowledge, perceptiveness and skilful supervision of most of my work. I wish to thank him dearly for this.

I am also indebted to my trusted colleague, Mr Mark Courtney, for numerous discussions on the more technical aspects of the thesis; to Messrs. Oakley West, Cartographer at the Geography Department of Rhodes University, and Mike Handley of the Printing Unit, for the extremely competent way in which they assisted with the diagrams; and to Miss Heather Kew who typed the thesis both efficiently and cheerfully. The financial assistance of the Human Sciences Research Council is gratefully acknowledged.

To my wife, Loekie, I am deeply thankful for the love and moral support she has given me during the time of writing this thesis.

Philip Black
Grahamstown
December 1981
# CONTENTS

<table>
<thead>
<tr>
<th>Acknowledgements</th>
<th>.................................................. (ii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>................................................................ (vi)</td>
</tr>
<tr>
<td>Introduction</td>
<td>................................................................ 1</td>
</tr>
<tr>
<td><strong>Chapter 1:</strong></td>
<td><strong>General Equilibrium and Economic Development:</strong></td>
</tr>
<tr>
<td></td>
<td><strong>A Neoclassical Approach</strong> ............................ 9</td>
</tr>
<tr>
<td>(1)</td>
<td>The Two-sector Model of General Equilibrium:</td>
</tr>
<tr>
<td></td>
<td>Basic Assumptions ....................................... 12</td>
</tr>
<tr>
<td>(2)</td>
<td>Static Equilibrium: Welfare Maximization and</td>
</tr>
<tr>
<td></td>
<td>Pareto Optimality ........................................ 14</td>
</tr>
<tr>
<td>(3)</td>
<td>The Comparative Statics of Economic Under-</td>
</tr>
<tr>
<td></td>
<td>development ................................................ 22</td>
</tr>
<tr>
<td>(4)</td>
<td>Comparative Static Equilibrium: Exit Under-</td>
</tr>
<tr>
<td></td>
<td>development ................................................ 26</td>
</tr>
<tr>
<td>(5)</td>
<td>Dynamic Equilibrium and the Steady State .......... 30</td>
</tr>
<tr>
<td>(6)</td>
<td>Conclusion: The Critical Assumptions .............. 35</td>
</tr>
<tr>
<td><strong>Chapter 2:</strong></td>
<td><strong>Economies of Scale, Technological Progress and</strong></td>
</tr>
<tr>
<td></td>
<td><strong>The Factor-proportions Problem</strong> .................. 38</td>
</tr>
<tr>
<td>(1)</td>
<td>Differential Returns to Scale: The Theory of</td>
</tr>
<tr>
<td></td>
<td>Cumulative Causation .................................... 39</td>
</tr>
<tr>
<td>(2)</td>
<td>Differential Returns and Dynamic Equilibrium .... 45</td>
</tr>
<tr>
<td>(3)</td>
<td>The Factor-proportions Problem ...................... 47</td>
</tr>
<tr>
<td>(4)</td>
<td>Inappropriate Technology and the Factor-</td>
</tr>
<tr>
<td></td>
<td>proportions Problem ..................................... 58</td>
</tr>
<tr>
<td>(5)</td>
<td>Policy Implications of the Factor-proportions</td>
</tr>
<tr>
<td></td>
<td>Problem .................................................... 62</td>
</tr>
<tr>
<td>(6)</td>
<td>Economic Growth and Fixed Factor Proportions:</td>
</tr>
<tr>
<td></td>
<td>Further Policy Considerations ..................... 66</td>
</tr>
</tbody>
</table>
## Chapter 3: Market Imperfections and Economic Underdevelopment: The Closed Economy

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Monopoly, Monopsony and Economic Underdevelopment</td>
<td>72</td>
</tr>
<tr>
<td>(2) Factor Market Imperfections and Economic Dualism</td>
<td>78</td>
</tr>
<tr>
<td>(3) Economic Dualism and Rural-Urban Migration</td>
<td>85</td>
</tr>
<tr>
<td>(4) The Urban Informal Sector</td>
<td>91</td>
</tr>
<tr>
<td>(5) General Policy Implications</td>
<td>96</td>
</tr>
</tbody>
</table>

## Chapter 4: Market Imperfections, International Trade and Policy

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Introduction: The Gains from Trade</td>
<td>102</td>
</tr>
<tr>
<td>(2) The Prebisch-Singer Thesis</td>
<td>108</td>
</tr>
<tr>
<td>(3) Tariff Protection and the Terms of Trade</td>
<td>114</td>
</tr>
<tr>
<td>(4) The Infant-industry Argument for Protection</td>
<td>120</td>
</tr>
<tr>
<td>(5) Market Imperfections and Trade Policy</td>
<td>123</td>
</tr>
</tbody>
</table>

## Chapter 5: Income Redistribution, Employment and Economic Growth

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Introduction</td>
<td>132</td>
</tr>
<tr>
<td>(2) Income Redistribution and Employment: The ILO Argument</td>
<td>133</td>
</tr>
<tr>
<td>(3) Income Redistribution and the Demonstration Effect</td>
<td>138</td>
</tr>
<tr>
<td>(4) Income Redistribution and Economic Growth</td>
<td>142</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>147</td>
</tr>
<tr>
<td>APPENDIX 1: The Social Welfare Function</td>
<td>152</td>
</tr>
<tr>
<td>APPENDIX 2: Welfare Maximization and Pareto Optimality</td>
<td>157</td>
</tr>
<tr>
<td>APPENDIX 3: Factor Mobility and the Neoclassical Steady-State</td>
<td>160</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>161</td>
</tr>
</tbody>
</table>
This thesis is mainly concerned with the question whether 'conventional' economic theory - especially the neoclassical theory of general equilibrium - is sufficiently flexible to accommodate the particular conditions prevailing in the developing countries. It is argued that most existing theories of economic underdevelopment adopt an interpretative approach which essentially amounts to relaxing some of the chief assumptions of the neoclassical theory. When applied to the two-sector model of general equilibrium, these theories generally yield predictions which are vastly different from those associated with the neoclassical assumptions of perfect competition, unlimited factor substitutability and unrestricted resource mobility.

Several theories seek to explain the development problem in terms of the specific production processes used in poor countries. Myrdal's (1957) theory of cumulative causation, for example, effectively introduces increasing returns to scale in at least one sector or region of the economy; in contrast to the neoclassical theory, he thus envisages a cumulative process of regional divergence in the output level per worker. Similarly, Richard Ecaus's (1955) explanation of the "factor-proportions problem" is based on the assumption of limited factor substitutability. This enables him to establish the existence of a so-called "unemployment equilibrium", thus implying that developing countries may
be faced with a conflict between the objective of maximizing social welfare on the one hand, and that of full employment on the other. More recently, Leibenstein (1960) has shown that this trade-off may be complicated by the introduction of capital-biased technological inventions and innovations. The solution to the factor-proportions problem consists in the adoption of more appropriate, usually labour-biased technologies, increased capital formation and a reduction in the rate of population growth.

Much of the postwar literature on economic development has focused on the imperfectly competitive structure of the product and the factor markets in developing countries. Myint (1954) has highlighted the role played by monopolies and oligopolies during the "opening-up" process of economic development. Likewise, both Lewis's (1954) dualist theory and Todaro's (1969; 1971) model of rural-urban migration attempt to explain the unemployment problem in terms of various factor price distortions. In an international context, Prebisch (1950; 1959) and Singer (1950) have again shown how prevailing differences in the structure of markets between developed and developing countries may turn the terms of trade against the latter; using a two-sector model, Bhagwati (1958) has demonstrated that such a deterioration in the terms of trade could bring about a net decrease in the welfare level of the countries concerned. Generally, the policy measures relevant to the "market imperfections" problem include the creation of job opportunities in the rural (rather than urban) sector, the encouragement of informal-sector
enterprises, and the imposition of factor taxes and subsidies as a means of counteracting the adverse effect of factor price distortions on employment.

A more recent approach to the unemployment problem is the plea by the International Labor Office (1970; 1972) for a redistribution of income within the developing countries. In terms of the two-sector model, such a policy may well succeed in eliminating labour unemployment caused by fixed factor proportions and/or factor price distortions. It should be realized, though, that a redistribution of income may lower the aggregate savings level, and hence also the growth rates of capital and labour employment in the economy.

On the whole, it would seem that these theories do indeed adopt a modified version of the neoclassical theory in providing a fairly comprehensive explanation of the economic problems of labour unemployment, low incomes and inequality.
"The major inadequacies of conventional economies are that the analysis focuses on the wrong factors, and the models do not fit at all closely the way in which non-industrial economies operate".

D. Seers (1963, p. 83)

"Over a wide range the relevance of economic analysis to poor countries is not in question, since some of the propositions of economics derive directly from the universal limitation of resources".

P.T. Bauer (1963, p. 360)

The views expressed in the above quotations are characteristic of two distinct schools in the field of development economics. On the one hand it is argued that economic theory in the 'conventional' sense is wholly or largely irrelevant to the problem of underdevelopment. Adherents to this view contend that received micro- and macroeconomic theory should either be changed extensively, or abandoned completely and replaced by a theoretical structure whose domain extends beyond the realm of the "economic", and whose methodology and basic premises would be more appropriate to actual conditions prevailing in the poor countries. Opposing views hold that conventional economic theory is either quite relevant to developed and developing countries alike, or that it could be made more applicable simply by modifying some of its assumptions.

According to Myrdal (1968; 1970; 1973), Seers (1963; 1971; 1971-72), Szentes (1976) and others,¹ the irrelevance of economic

---

¹ See, for example, Sweezy (1970), Streeton (1972, ch. 5) and Lipton (1977, ch. 4).
theory derives from the fact that its assumptions are generally drawn from the special conditions existing in the industrially advanced countries. Most economists are insufficiently trained to recognize, let alone analyze, the social and political factors determining economic progress in the developing countries. While such an "omission" may be permissible in the context of the advanced economies, the relative importance of these "non-economic" variables in developing countries renders a purely economic approach largely meaningless. Myrdal (1973) and Seers (1963) in particular question the methodology and "realism" embodied in several theoretical approaches to the study of economic under-development, such as the neoclassical theory of resource allocation, Keynes's multiplier principle, growth and "stages of growth" theories developed by Harrod (1939), Domar (1946) and Rostow (1956; 1960). These theories are said to rely on patently unrealistic assumptions and generally fail to come to grips with the essentially dynamic task of promoting economic development in poor countries. In brief, adopting the conventional approach introduces a methodological and ideological bias which tends to "distort and prejudice our view into a particular direction".2

Although few economists would deny that these criticisms are at least partly justified, most would agree that they hardly constitute an argument for the abandoning of received economic theory. Apart from the difficulties involved in defining and quantifying such "non-economic" variables as "... the social

structure; political forces, attitudes and institutions", their incorporation into models designed to explain development problems is likely to produce excessively complex propositions and policy prescriptions. Moreover, it seems premature to advocate the abandonment of conventional economic theory without first establishing whether the existing theories could be modified in such a way as to accommodate the various criticisms listed above. In a similar vein, Myint (1967) has argued that the static theory of resource allocation associated with the neoclassical school may well assume greater significance in the context of a more comprehensive, dynamic approach to the study of economic underdevelopment. This is because the growth potential of a country, whether it is developed or developing, depends at least partly on whether it is able to utilize its given resources in a socially efficient way: "a country's absorptive capacity must to a large extent depend on its ability to avoid serious misallocation of resources".

Myint (1954; 1967), Bauer (1963; 1971) and others maintain that existing economic theory is sufficiently adaptable to fit the particular conditions of the developing countries. This proposition is similar to what Chenery (1975, p. 310) once called the "structuralist approach", according to which an attempt is made to "identify specific rigidities, lags and other characteristics of the structure of developing countries that affect economic adjustments and the choice of development policy". Although

5. See Schultze (1973) and Chenery (1975).
elements of "structuralism" are present in several postwar investigations into problems of development - including Myrdal's (1957) notion of cumulative causation, Lewis's (1954) labour-surplus economy, Prebisch's (1950; 1959) study of the international terms of trade, and recent work on the urban informal sector - it seems rather odd that many of these analyses have been viewed as representing a "radically new" or "revolutionary" approach to the study of economic underdevelopment. The fact of the matter would seem that most "structuralists" usually employ conventional tools of economic analysis which can be easily fitted into a classical, neoclassical or Keynesian theoretical framework; far from breaking down conventional economics, the structuralist approach may yet be seen one day as having reinforced and enriched the existing body of economic theory.

It is within this general context that Myint (1967, p. 117) has stressed the need for ".... a general practitioner to act as a middleman between different specialized fields of development economics and also between development economics and general economics ..... His aim should be to try to apply the existing economic theory in a more realistic and fruitful way to suit the varying conditions of different types of underdeveloped country".

6. See Kay's (1975) and Leys's (1977) criticisms of the dependency thesis.
7. One is reminded of Kay's (1975, p. 104) scathing remark that the dependency thesis ".... is an eclectic combination of orthodox economic theory and revolutionary phraseology".
What, then, should be the task of such a "general practitioner"? Should he try to integrate the various 'structuralist hypotheses' into a general theory of economic underdevelopment? Or should he adopt a partial approach in which the distinct characteristics of the developing economies are analyzed independently within a suitably modified variant of one or more of the existing theories? The answer to such a question may lie in the fact that the size and nature of the development problem tend to vary greatly within and between different developing countries. Such diversity severely limits the number of generalizations that can be effectively applied to all developing countries at the same time. Even if it were possible to construct a 'general' theory of development, its inevitably complex structure is likely to yield predictions of very limited practical value; in Friedman's (1966, p. 14) words, such a 'general' theory would explain "little by much".

The most appropriate course of action would seem to be the development of "partial" theories as a means of analysing specific economic problems common to most or ideally all developing countries. In this context it is worth noting Leibenstein's (1966) suggestion that the "efficiency" of the partial approach depends *inter alia* on whether it can be applied within a suitable "analytical framework". The aim of such a framework should be to simplify and reduce the multitude of real-world observations to "a small enough bundle of general concepts that they may be discussed efficiently"; it should be looked upon as ".... the mold out of which specific types of theories are made".8

One of the most useful of such existing frameworks is arguably represented by the neoclassical theory of general equilibrium. This view is not necessarily based on a belief that the underlying assumptions of the theory are in any sense "realistic", or that its behavioural relationships are always capable of yielding accurate predictions. Rather, the usefulness of general equilibrium theory derives from its capacity to accommodate a large variety of alternative assumptions. It is this built-in flexibility that enables the theory to yield alternative predictions applicable to many different real-world situations.

Generally, there can be little doubt that economic conditions in the developing countries do differ vastly from those prevailing in the industrially advanced countries. What is at stake is the question whether conventional economic theory - especially the neoclassical theory of general equilibrium - could be adapted in such a way as to allow for the particular conditions of the developing countries. It is with this question that the present thesis is mainly concerned. We begin with an introductory chapter on the neoclassical theory of general equilibrium. The main purpose of this chapter is to provide an appropriate analytical framework within which the various "partial" theories of underdevelopment may be subsequently discussed in a meaningful way. While the two-sector model of general equilibrium will be used throughout the thesis, chapters 2 to 5 attempt to show that existing theories have in common an analytical approach which

9. The theory does not, of course, lay claim to any such properties. (See Kaldor (1972, pp. 1237-1238); Friedman (1966)).
essentially amounts to relaxing some of the chief assumptions of the neoclassical theory. Specifically, Chapter 2 considers several theories which seek to explain the development problem in terms of particular production processes in the poor countries: for example, while Myrdal's (1957) theory of cumulative causation effectively introduces differential returns to scale into the two-sector model, Richard Eckaus's (1955) analysis of the "factor-proportions problem" is again based on the "alternative" assumption of limited factor substitutability. Likewise, Leibenstein (1960) has shown that the factor-proportions problem may be aggravated by the introduction of inappropriate technological inventions and innovations.

Chapters 3 and 4 are both concerned with the fact that the product and factor markets in most developing countries are generally characterized by imperfect competition. Chapter 3 discusses Myint's (1954) view of the role played by foreign-owned monopolies during the so-called "opening-up" stage of economic development. Also, an attempt is made to place Lewis's (1954) dualist theory, Todaro's (1969, 1971) model of rural-urban migration and the phenomenon of the urban informal sector within the general context of factor market imperfections. Chapter 4 examines the effect of various price distortions on the static and dynamic gains derived from international trade. In particular, the chapter considers Prebisch's (1950, 1959) and Singer's (1950) contention that the differences in market structure between developed and developing countries have turned the terms of trade
against the latter. This is followed by an assessment of the relative efficacy of tariff protection and domestic taxes and subsidies in eliminating international trade disequilibria arising from the existence of domestic price distortions.

Chapter 5 considers the recent debate on the employment potential associated with a policy of income redistribution in favour of low-income groups within the developing countries. It is shown there, that while such a policy may well give rise to a net increase in labour employment in the short run, this is likely to be eventually offset by changes in individual tastes resulting from a so-called demonstration effect. Similarly, available evidence indicates that a redistribution of income may cause a decrease in the aggregate savings level, and hence also in the growth rates of capital, employment and output. The thesis concludes with a brief summary of its main findings and relevant policy implications.
CHAPTER 1

GENERAL EQUILIBRIUM AND ECONOMIC DEVELOPMENT:
A NEOCLASSICAL APPROACH

Although there are good reasons why the ceteris paribus assumption has been widely used in both micro- and macroeconomic analysis, this should not cloud the fact that, if carried too far, "holding other things constant" may yield inadequate, if not misleading, results. In the field of growth economics, for example, the predictive value of the Harrod-Domar and neoclassical models is limited by their sole reliance on demand and supply factors, respectively. Similarly, international trade theorists have either ignored demand factors altogether or, at least, impounded them in a ceteris paribus assumption. Likewise, it is probably fair to say that the economist's rather belated "discovery" of the stagflation phenomenon is due in part to his having neglected changing supply conditions; and indeed, only time will tell whether the recent "supply-side" economics may prove to be similarly constrained. However useful the ceteris paribus assumption may be, there can be no doubt that its painstaking application in economic analysis has hardly done justice to those who created and perfected the "general cases" in the first place.

In the field of development economics the same criticism seems to apply only to a relatively limited extent; in point of fact, the failings and frailties of development economics partly stem from an unwillingness or inability to hypothesize on the basis of ceteris paribus. Most "theories" of economic underdevelopment either take too narrow a view of the problem at hand, or merely recount real
world complexities in a rather equivocal and often haphazard manner. To the former category belong *inter alia* several growth models\(^1\) and strategies,\(^2\) demographic studies\(^3\) and dualist theories of economic development.\(^4\) Other writers have again tried to account for too many variables at once, thus limiting the testability and general predictability of their respective analyses: it is perhaps paradoxical that none of the recent "interpretations",\(^5\) "dramas",\(^6\) "non-communist manifesto"\(^7\) and "crises"\(^8\) have even come close to providing a comprehensive and generally valid explanation of economic underdevelopment. While it cannot be denied that these studies have provided useful insights into the nature of the development problem itself, it seems nevertheless either premature to claim discovery of a "main hypothesis...a vision of the general theory which we are all yearning for,"\(^9\) or it may be simply futile to even "yearn" for such a theory.

In the remainder of this thesis we shall attempt to steer a middle course by adopting a simplified version of the Walrasian theory of general economic equilibrium. Specifically, our analysis will be limited to the familiar two-sector model, in which two consumers are


2. We refer to the literature on balanced versus unbalanced growth: Rosenstein-Rodan (1943), Hirschman (1958), Scitovsky (1954 and 1959), Streeten (1959), Nurske (1953) and Lewis (1955).


assumed to supply two production factors to firms producing two commodities for final demand purposes,\textsuperscript{10} it is perhaps worth noting too that the symbols X and Y will be used alternately to denote commodities, firms, sectors and regions. Although the two-sector model is not without its shortcomings, it does at least allow one to consider supply and demand conditions simultaneously, both from a static and a dynamic perspective. Similarly, it is generally agreed that the chief conclusions of the two-sector model are valid also for an economy containing several consumption goods and capital goods sectors.\textsuperscript{11}

This chapter falls into six parts: section 1 provides a brief summary of the basic assumptions underlying the two-sector model, which are then used in section 2 to derive the various conditions necessary to achieve a Pareto-optimal allocation of resources; section 3 introduces appropriate exogenous changes into the two-sector model, in an attempt to explain the emergence of economic underdevelopment from a purely neoclassical perspective; sections 4 and 5 discuss the comparative static and dynamic adjustment mechanisms operating under conditions of perfect competition; and finally, section 6 assesses the extent to which the chief assumptions of the theory approximate real-world conditions in developing countries.

\textsuperscript{10} See, for example, Johnson (1971), Krauss and Johnson (1974) Simpson (1975) and Baldry (1980).

\textsuperscript{11} Note, for example, Meade's (1961, p. x) admittedly cautious view: "...I have a strong hunch that the main result would be not very substantially to alter the basic conclusions of the present (two-sector) analysis, but very greatly to increase the possibility of substitution between the various factors of production".
1. The Two-sector Model of General Equilibrium: Basic Assumptions.

To begin with, it seems worth giving a brief account of the basic assumptions of the two-sector model, especially since many of these will be subsequently relaxed in the chapters to follow. The most important ones are listed below:

(i) There are two production factors, capital (K) and labour (L) both of which are infinitely divisible and owned by two individual consumers, A and B. The total supply of K and L as well as its distribution among A and B are exogenously determined.

(ii) Each consumer derives satisfaction or 'utility' from consuming two substitutable commodities, X and Y. Consumer tastes are determined independently and exogenously, which rules out external economies and diseconomies in consumption. Preferences are transitive while each consumer is assumed to prefer more of both goods. These assumptions imply that consumer preferences may be represented by a series of indifference curves that are continuous, non-intersecting and convex to the origin.

(iii) Each commodity is produced by one firm only. Although the two production functions are non-identical, both are characterized by unlimited factor substitutability and diminishing marginal productivities, thus giving rise to smooth isoquants that are convex to the origin. The further assumption that each production function is also subject to constant returns to scale, rules out internal and external (dis)economies in production.
(iv) The sole objective of each consumer and firm is maximization of utility and profit, respectively.

(v) Both the factor markets and the commodity markets are perfectly competitive, which implies the familiar assumptions of homogeneity, perfect knowledge and the presence of large numbers of small-sized buyers and sellers; in addition, we assume that all resources are completely mobile, both in the occupational and spatial sense of the word.

(vi) There exists a social welfare of the form

\[ W = W(U_a, U_b) \] (1.1)

where \( W \) represents the level of social welfare, and \( U_a \) and \( U_b \) are the utility levels of A and B respectively. Since the chief characteristics of the social welfare function are discussed in Appendix 1, suffice it to mention here that it presupposes an ethical valuation of the relative worthiness of A and B.

We might finally note that assumptions (ii) to (v) ensure that a general equilibrium does in fact exist.\(^{12}\) Similarly, if we rule out the possible existence of an upward-sloping demand curve, then it can be shown that the same assumptions also guarantee the existence of a stable and unique equilibrium.\(^{13}\)

---


The welfare function introduced in the previous section may be viewed as the overall objective of society, which needs to be maximized subject to a set of given constraints. These constraints consist of (i) the fixed endowments of K and L; (ii) the tastes of A and B which are represented here by the respective individual utility functions; and (iii) the state of the technology as indicated by the production functions for X and Y. The maximization problem may thus be formulated as follows: 14

Maximize \( W = W(U_a, U_b) \) \hspace{1cm} (1.1)

subject to

\[ U_a = U_a(Q^a_x, Q^b_x) \] \hspace{1cm} (1.2)
\[ U_b = U_b(Q^b_x, Q^b_y) \]

\[ Q^a_x + Q^b_y = Q_x = Q_x(K_x, L_x) \] \hspace{1cm} (1.3)
\[ Q^a_y + Q^b_y = Q_y = Q_y(K_y, L_y) \]

and

\[ \bar{K} = K_x + K_y \] \hspace{1cm} (1.4)
\[ \bar{L} = L_x + L_y \]

where \( Q^a_x, Q^a_y \) and \( Q^b_x, Q^b_y \) are the quantities of X and Y consumed by A and B respectively; \( K_x, L_x \) and \( K_y, L_y \) are the quantities of K and L used in the production of X and Y, respectively; and \( \bar{K} \) and \( \bar{L} \) are the respective endowments of K and L.

14. Similar procedures are followed by Layard and Walters (1978, ch. 1) and Henderson and Quandt (1971, ch. 4).
The solution to this maximization problem is derived fully in Appendix 2. It is worth noting here, however, that the solution entails all the familiar conditions for a Pareto-optimal allocation of resources. These are reproduced below:

(i) The first condition refers to production efficiency in the economy as a whole; or

\[
\frac{\partial Q_x / \partial L_x}{\partial Q_x / \partial K_x} = \frac{w}{r} = \frac{\partial Q_y / \partial L_y}{\partial Q_y / \partial K_y}
\]  

(1.5)

where \( w \) and \( r \) are the prices of labour and capital respectively. Condition (1.5) represents equality between the marginal rates of technical substitution of \( L \) for \( K \) in the production of \( X \) and \( Y \), respectively; or

\[
\text{MRTS}_{L,K}^X = \text{MRTS}_{L,K}^Y
\]  

(1.5')

This condition may be illustrated with the aid of Figure 1.1. The dimensions of the box diagram shown in Figure 1.1(a) are determined by the given supplies of \( K \) and \( L \): the vertical axes depict units of \( K \) and the horizontal axes units of \( L \), while \( O_x \) and \( O_y \) are the origins for commodities \( X \) and \( Y \) respectively. The contract curve, \( O_x O_y \), represents the locus of input ratios for which the respective marginal rates of technical substitution are equal; that is, each point along \( O_x O_y \) designates the maximum attainable quantity of one commodity, given the quantity of the other. Consequently, since each input ratio along \( O_x O_y \) corresponds to a unique output combination along the production possibility or transformation curve, \( T T' \) in Figure 1.1(b), it follows that equality between the individual marginal rates of technical substitution is a necessary condition for output maximization in the economy as a whole.
FIGURE 1.1
(ii) The second optimality condition implies that

\[
\frac{\frac{\partial U_a}{\partial q_{ax}}}{\frac{\partial U_a}{\partial q_{ay}}} = \frac{P_x}{P_y} = \frac{\frac{\partial U_b}{\partial q_{bx}}}{\frac{\partial U_b}{\partial q_{by}}}
\]

(1.6)

where \(P_x\) and \(P_y\) represent the prices of commodities X and Y respectively. Condition (1.6) is the familiar equality between A and B's respective marginal rates of substitution of X and Y; or

\[
\text{MRS}^a_{x,y} = \text{MRS}^b_{x,y}
\]

(1.6')

This is illustrated in Figure 1.2 below, where the (arbitrarily chosen) output mix \(O_b\) on the transformation curve determines the dimensions of the corresponding box diagram for exchange: the origins for individuals A and B are given by \(O_a\) and \(O_b\) respectively. The exchange contract curve, \(O_aO_b\), represents the locus of output mixes for which the respective individual marginal rates of substitution are equal; in other words, each point along \(O_aO_b\) indicates the maximum attainable utility of one consumer, given the utility level of the other. It thus follows that equality between the marginal rates of substitution is a necessary condition for maximization of the combined level of individual utilities.\(^{15}\)

---

15. This is simply another way of defining our second condition for Pareto optimality; it does, of course, imply that individual utilities are comparable and measurable.
FIGURE 1.2
(iii) The third or "top level" condition may be derived from the following identity,\(^\text{16}\)

\[ V = P_x Q_x + P_y Q_y \]  

where \( V \) is the value of the total or combined product of \( X \) and \( Y \). Solving for \( Q_y \) in (1.7) gives:

\[ Q_y = V/P_y - (P_x/P_y) Q_x \]  

which is the mathematical equivalent of an isorevenue or commodity price line.\(^\text{17}\) Two such price lines are shown in Figure 1.2, viz. the parallel lines labelled \( mm' \) and \( nn' \). The fact that \( nn' \) represents a higher \( V \) then \( mm' \), means that the value of total output is in fact maximized at \( O_b \), where \( nn' \) is tangent to the transformation curve, \( TT' \); that is,

\[ \frac{dQ_y}{dQ_x} = \frac{P_x}{P_y} \]  

where \( \frac{dQ_y}{dQ_x} \) is the slope of the transformation curve, or the marginal rate of product transformation between \( X \) and \( Y \) (MRPT\(_{x,y}\)). Accordingly, a necessary condition for maximization of the value of total output is equality between the MRPT\(_{x,y}\) and the corresponding commodity price ratio.

Similarly, combining (1.6), (1.6') and (1.9) gives

\[ MRS_{a_{x,y}} = MRS_{b_{x,y}} = \text{MRPT}_{x,y} \]

representing equality between the rate at which consumers are willing to substitute \( X \) for \( Y \) at the margin, and the rate at

\(^{16}\) An alternative - and more conventional - derivation is provided in Appendix 2.

\(^{17}\) See, for example, Koutsoyiannis (1979, pp. 535-536).
which it is technically possible to do so. This condition is indicated in Figure 1.2 by the parallel price lines $mm'$ and $nn'$, passing through the equilibrium output ratios $O_b$ and $E'$, respectively.

A closely related condition, sometimes referred to as the "social justice condition", is based on the following equality,

$$\frac{aW / aQ_x}{aW / aQ_y} = \frac{P_x}{P_y}$$

(1.11)

where the lefthand side may be regarded as the rate of change of a community indifference curve in output space, or the community marginal rate of substitution of $X$ for $Y$ ($CMRS_{x,y}$). Although the related concepts of the social welfare function and the community indifference curve are discussed in Appendices 1 and 2, it should be emphasised here that the latter is derived on the assumption that the two commodities, $X$ and $Y$, are optimally distributed between the two consumers, A and B.

Combining (1.10) and (1.11) we get:

$$CMRS_{x,y} = MRS^a_{x,y} = MRS^b_{x,y} = MRPT_{x,y}$$

(1.12)

which brings together all the necessary and sufficient conditions for the maximization of social welfare — the "optimum optimorum". This is shown as the output mix $E$ in Figure 1.3 below, where the transformation curve is tangent to the highest attainable community indifference curve, $W_2$. Similarly, at $E$ and $E'$ in Figure 1.3 the social and individual welfare levels are all maximized simultaneously, subject to the given set of constraints.

19. See Appendix 2.
FIGURE 1.3
We may therefore conclude that Pareto optimality per se is only a necessary condition for the maximization of social welfare. The sufficient condition includes (i) equality between the community marginal rate of substitution and the marginal rate of product transformation (e.g. point E in Figure 1.3), and (ii) equality between the latter rates and the individual marginal rates of substitution of X for Y (at point E'). These various rates of substitution are, of course, brought into equality via the price mechanism of a perfectly competitive economy; and in Figure 1.3 the relevant ratio of $P_x$ to $P_y$ is given by the slope of the parallel price lines mm' and nn'.

3. The Comparative Statics of Economic Underdevelopment

What is the significance of Pareto optimality and the competitive equilibrium for the process of economic development? Here it is necessary to adopt both a comparative static and a dynamic approach to the general equilibrium model outlined in the previous section. In particular, we need to know how the system evolves from one state to another over time — for example, from an initial state of relative poverty (or affluence) to one of relative affluence (or poverty); and for this purpose, it is convenient first to define the development problem formally by introducing appropriate exogenous changes into our two-sector model of general equilibrium.

22. Or more precisely, by the omniscient Walrasian auctioneer.
To begin with, consider the following suggested scenario of the emergence of economic underdevelopment. Imagine an economy which has led a relatively isolated existence for a considerable period of time: in a static sense, it is enjoying the kind of "optimum optimorum" discussed in the previous section, while dynamically it is approximating a so-called steady state in which output, capital and labour are all growing at more or less the same constant rate. Suppose the economy is now "opened up" exogenously by an inflow of capital and/or labour, the effect of which is either to raise or lower the prevailing capital/labour ratio. It seems reasonable also to assume that the opening-up process is localized and confined to one productive sector only, assuming that each of the goods (X and Y) is also (spatially) produced in a specific region within the economy. The opening-up process accordingly widens existing regional differences in the capital/labour ratio, creating similar discrepancies in the marginal productivities and output levels per worker within the economy as a whole.

In order to illustrate the emergence of underdevelopment geometrically, we thus distinguish between two regions, X and Y, each of which produces the correspondingly lettered commodity. The initial (static) equilibrium occurs at the output ratio \( E \) in Figure 1.4(b), which corresponds to the input ratio \( E' \) in Figure 1.4(a). These ratios represent production of the output levels \( X_2 \) and \( Y_2 \) respectively, at uniform factor prices given by

23. Although this term is borrowed from Myint (1954), we attach a much narrower meaning to it than did Myint. In either case, however, the "opening-up process" may be prompted by economic, political and/or military considerations.
FIGURE 1.4
the slope of the factor price line, aa', passing through E'.

Suppose now that the quantity of capital in region Y increases by \( y_0y' \) units in Figure 1.4(a). This causes the contract curve to shift from \( 0_0x_0y \) to \( 0_0x_0y' \) in Figure 1.4(a), which implies in turn a shift in the transformation curve from \( TT' \) to \( SS' \) in Figure 1.4(b).²⁴

Now, the effect on factor prices depends in part on the assumption of constant returns to scale; that is, the fact that the marginal productivities of capital and labour are constant proportionate functions of the relevant capital/labour ratios. It can be shown, for example, that the partial derivatives of (1.3) are:

\[
\frac{\partial Q_x}{\partial K_x} = \frac{a}{K_x^*}; \quad \frac{\partial Q_x}{\partial L_x} = bK_x^*
\]

and

\[
\frac{\partial Q_y}{\partial K_y} = \frac{e}{K_y^*}; \quad \frac{\partial Q_y}{\partial L_y} = fK_y^*
\]

where \( K_x^* \) and \( K_y^* \) represent the capital/labour ratios used in the production of commodities X and Y, respectively; and \( a, b, e \) and \( f \) are positive constants. According to (1.13) a given increase in \( K_y^* \) will ceteris paribus raise the marginal product of region Y's labour force, and lower its marginal product of capital, relative to those in region X. These changes in the marginal productivities of region Y are shown in Figure 1.4(a) by the difference between the slope of the new factor price line, bb', through E', and the slope of the original price line, aa'. Although the equilibrium input ratio (temporarily) remains at E', it is clear that region Y is now using more capital

²⁴ On the nature of these shifts, see Rybczynski (1955).
relative to labour in producing the higher output level given by $Y_3'$. The latter increase in the output of $Y$ causes a decrease in its relative price — indicated in Figure 1.4(b) by the difference between the slope of the commodity price line, $pp'$, and the slope of the original one, $nn'$ passing through $E$. Similarly, the equilibrium output ratio has moved from $E$ to $F$ in Figure 1.4(b), where the community indifference curve labelled $W_3$ is tangent to the new commodity price line, $pp'$. It is important to note, however, that $E'$ and $F$ now represent sub-optimal allocations with respect to $O_xO_y'$ and SS' in Figures 1.4(a) and 1.4(b), respectively.

4. Comparative Static Equilibrium: Exit Underdevelopment

Is the absence of Pareto optimality a permanent or temporary state of affairs? The answer depends, of course, on the assumptions of perfect competition, unlimited substitutability and unrestricted factor mobility, which together ensure that any deviation from equilibrium will be "... automatically redressed by the incentive-providing mechanisms of the market".\textsuperscript{25} The equilibrating nature of the neoclassical economy may be explained by a simple extension of the two-sector model. Consider, for example, the following differentials of equations (1.3):

$$dQ_x = \left(\frac{aQ_x}{aK_x}\right) dK_x + \left(\frac{aQ_x}{aL_x}\right) dL_x$$

and

$$dQ_y = \left(\frac{aQ_y}{aK_y}\right) dK_y + \left(\frac{aQ_y}{aL_y}\right) dL_y$$

\textsuperscript{25.} Fei and Ranis (1964). See also Ohlin (1967, Part 3).
Since we are interested here in the interregional allocation of a given quantity of resources, let us temporarily assume that savings and the natural growth of population are both equal to zero. Similarly, let the spatial mobility of capital and labour be proportionately related to the corresponding differences in the marginal productivities; that is,

\[ dK_x = -dK_y \]
\[ = g\left(\frac{aQ_x}{aK_x} - \frac{aQ_y}{aK_y}\right) \]  

(1.15)\hfill (1.16)

or, substituting (1.13) into (1.16),

\[ dK_x = g\left(\frac{a}{K_x^*} - \frac{e}{K_y^*}\right) \]  

(1.16')

And similarly,

\[ dL_x = dL_y \]
\[ = h\left(\frac{b}{K_x^*} - \frac{f}{K_y^*}\right) \]  

(1.17)\hfill (1.18)

where \( g \) and \( h \) are positive constants. Substituting (1.15) to (1.18) into (1.14) gives.

\[ dQ_x = \left(\frac{a}{K_x^*}\right) g\left(\frac{a}{K_x^*} - \frac{e}{K_y^*}\right) + \left(\frac{b}{K_x^*}\right) h\left(\frac{b}{K_x^*} - \frac{f}{K_y^*}\right) \]  

(1.14')

and

\[ dQ_y = -\left(\frac{e}{K_y^*}\right) g\left(\frac{a}{K_x^*} - \frac{e}{K_y^*}\right) - \left(\frac{f}{K_y^*}\right) h\left(\frac{b}{K_x^*} - \frac{f}{K_y^*}\right) \]

Now, since \( \left(\frac{a}{K_x^*}\right) > \left(\frac{e}{K_y^*}\right) \) and \( \left(\frac{b}{K_x^*}\right) < \left(\frac{f}{K_y^*}\right) \) by assumption, it follows from (1.14') that (i) labour will migrate from region \( X \) to region \( Y \) and capital in the opposite direction; (ii) these factor flows will in turn reduce the differences in the capital/labour ratio and marginal productivities; until (iii) a stable equilibrium is again reached when \( \left(\frac{a}{K_x^*}\right) = \left(\frac{e}{K_y^*}\right), \left(\frac{b}{K_x^*}\right) = \left(\frac{f}{K_y^*}\right) \)
and \( dQ_x = dQ_y = 0 \). The fact that \( K_x^* \) is now higher than before, while \( K_y^* \) is again lower, implies that there has been a convergence in the regional output level per labourer.\(^{26}\)

All this may be illustrated with the aid of Figure 1.5 below, which is a straight reproduction of Figure 1.4. At the input ratio \( E' \) in Figure 1.5(a), for example, the (real) price of capital is assumed to be higher in region \( X \), while the price of labour is again higher in region \( Y \). Accordingly, labour may be expected to flow from region \( X \) to region \( Y \) and capital in the opposite direction, thus raising (lowering) the capital/labour ratio and wage rate, and lowering (raising) the interest rate in region \( X(Y) \). These factor movements will continue until the interregional differences in factor prices have been entirely eliminated at, say, the input ratio \( G' \) on the contract curve, \( 0_{x0y} \) — indicating production of the output levels \( X_3 \) and \( Y_4' \) at uniform factor prices given by the slope of the new factor price line, \( cc' \) passing through \( G' \). Similarly, since production in each region is now higher than before, the economy is able to produce the preferred output combination \( G \) on the transformation curve \( SS' \) in Figure 1.5(b), representing the maximum attainable welfare level of \( W_4 \).

Generally, the above illustration serves to highlight the important role played by interregional factor movements in the equilization of factor prices. The equilibrating nature of such factor flows derives from the fact that production factors migrate from regions where their marginal

\(^{26}\) This follows from the Cobb-Douglas assumption that output per worker is a direct proportionate function of the capital/labour ratio.
FIGURE 1.5
productivity is low, to regions where it is high, thus causing a net increase in the (real) output of the economy as a whole: this is easily proved by setting

\[ dQ_t = dQ_x + dQ_y \]  

(1.19)

where \( dQ_t \) represents the change in the total or combined regional output. Substituting (1.14') into (1.19) and rearranging terms, we get

\[ dQ_t = g(a/K_x^* - e/K_y^*)^2 + h(bK_x^* - fK_y^*)^2 \]  

(1.19')

which shows that if

\[ (a/K_x^*) \geq (e/K_y^*) \quad \text{and/or} \quad (bK_x^*) \geq (fK_y^*) \]

then \( dQ_t > 0 \).

5. Dynamic Equilibrium and the Steady State

Turning to the dynamic equilibrium of the two-sector model, it becomes necessary to relax the assumptions of zero savings and a constant supply of labour. Both savings and the natural growth of labour are important sources of regional economic growth generally, and of regional factor mobility in particular; and it is our purpose here to determine whether and to what extent such factor movements affect the dynamic equilibrium of the perfectly competitive economy.

Before this is done, however, recall that our earlier economy was assumed to be spatially differentiated only with respect to the technical conditions of production; specifically,
capital's share of output was higher in region Y than in region X. Suppose now that the respective factor prices are initially the same everywhere, while the savings propensity and the natural growth rate of labour are both constant and spatially uniform. Under these assumptions, the growth rates of output in the two regions are given by the following familiar equations:

\[ q_x = \alpha \sigma_x + (1 - \alpha) n \]

and

\[ q_y = \beta \sigma_y + (1 - \beta) n \]  

where \( s \) and \( n \) represent the marginal savings propensity and the natural growth rate of labour, respectively; \( \sigma_x \) and \( \sigma_y \) are the (variable) output/capital ratios in regions X and Y; \( \alpha \) and \( \beta \) are capital's share of output in regions X and Y respectively; and \( \alpha < \beta \) by assumption. Now, the steady state is generally characterized by the fact that output and capital grow at the same constant rate; that is, \( q_x = s \sigma_x \) and \( q_y = s \sigma_y \), whence it follows from (1.20) that

\[ q_x = q_y = n \]  

Equation (1.21) simply states that in the absence of factor mobility, each region will experience steady-state growth independently, being determined as it is by the given growth rate of its labour supply. But since the latter growth rate

---

27. See, for example, Solow (1956), Swan (1956) and Meade (1961).

28. Both are assumed to be exogeneously determined by non-economic factors (See Meade (1961, p. 19)).
is assumed to be spatially uniform, it follows that both regions will grow at the same steady-state rate, \( n \); and similarly, the growth rate of the total or combined regional output will also equal \( n \). Likewise, the growth rate of output per worker will be zero everywhere.

Let us now restate in dynamic terms our earlier account of the opening-up process. At a given point in time, region \( Y \) experiences a (net) inflow of capital from abroad, which continues for a considerable period of time. This increase in the capital stock of region \( Y \), \textit{ceteris paribus}, raises the marginal productivity or wage rate of its labour force, and lowers its marginal product of capital, relative to those in region \( X \). If labour now starts to migrate from region \( X \) to region \( Y \), and capital in the opposite direction, then the growth rates of regional output become:

\[
q_x' = \alpha k_x + (1 - \alpha) l_x \\
q_y' = \beta k_y + (1 - \beta) l_y
\]

where \( k_x \) and \( k_y \) are the growth rates of capital in regions \( X \) and \( Y \) respectively; and \( l_x \) and \( l_y \) are the corresponding growth rates of the labour supply. Each of these factor growth rates is determined both by endogenous and exogeneous variables; for example,

\[
k_x = s_{\sigma x} + K_{yx} \\
l_x = n + l_{yx} \\
k_y = s_{\sigma y} + k_{fy} + k_{xy} \\
l_y = n + l_{xy}
\]
where \( k_{yx} \) \((k_{xy})\) is the rate at which capital is being transferred from region \( Y(X) \) to region \( X(Y) \); \( l_{yx} \) \((l_{xy})\) is the rate of labour migration from region \( Y(X) \) to region \( X(Y) \); and \( k_{fy} \) is the constant rate of foreign capital inflow into region \( Y \). Similarly, we know from (1.15) through (1.18) that:

\[
\begin{align*}
k_{yx} &= g \left( \frac{a}{K_x} - \frac{e}{K_y} \right) / K_x \\
k_{xy} &= -g \left( \frac{a}{K_x} - \frac{e}{K_y} \right) / K_y \\
l_{yx} &= h \left( \frac{bK_x^* - fK_y^*}{L_x} \right) \\
l_{xy} &= -h \left( \frac{bK_x^* - fK_y^*}{L_y} \right)
\end{align*}
\]

where \( K_x \) and \( K_y \) are the quantities of capital, and \( L_x \) and \( L_y \) the quantities of labour used in the production of \( X \) and \( Y \), respectively. Accordingly, substituting (1.24) into (1.23) into (1.22), we have

\[
\begin{align*}
q_x' &= \alpha \{ s \sigma_X + g \left( \frac{a}{K_x^*} - \frac{e}{K_y^*} \right) / K_x \} \\
&\quad + (1 - \alpha) \{ n + h \left( \frac{bK_x^* - fK_y^*}{L_x^*} \right) \} \\
q_y' &= \beta \{ s \sigma_Y - g \left( \frac{a}{K_x^*} - \frac{e}{K_y^*} \right) / K_y \} \\
&\quad + (1 - \beta) \{ n - h \left( \frac{bK_x^* - fK_y^*}{L_y^*} \right) \}
\end{align*}
\]  

(1.22')

Now, given that \((a/K_x^*) > (e/K_y^*)\) and \((bK_x^*) < (fK_y^*)\), there are two effects that can be distinguished on the basis of equations (1.22'). The first effect occurs more or less instantaneously and may be referred to as the "non-steady state" effect; the second is again a long term effect in the sense that it facilitates, and indeed accelerates, the movement toward steady-state growth in each region.
Firstly, given the presumed differences in factor prices, labour will migrate from region X to region Y and capital from region Y to region X; that is $k_{yx} > 0$, $k_{xy} < 0$, $l_{yx} < 0$ and $l_{xy} > 0$. The immediate or non-steady state effect is thus:

$$(q'_x - l_x) > 0$$

and

$$(q'_y - l_y) < 0$$

which indicates simply that the growth rate of output per worker will be positive in region X, but negative in region Y, thus implying a process of convergence in the regional output level per worker.

Secondly, the above factor flows will, ceteris paribus, reduce the differences in the capital/labour ratio and marginal productivities between regions X and Y; that is, they will lower (raise) $k_{yx}$($k_{xy}$) and raise (lower) $l_{xy}$($l_{yx}$). But these changes in the marginal productivities will be continuously offset by the successive inflows of foreign capital into region Y; and to the extent that they are, labour (capital) will continue to migrate from region X(Y) to region Y(X). Such continuous factor flows will, of course, reduce $\sigma_x$ and domestic savings in region X, while raising them again in region Y, until eventually each region has entered the steady-state when $q'_x = k_x$ and $q'_y = k_y$; that is,

$$q'_x = n + l_{yx}$$

and

$$q'_y = n + l_{xy}$$

(1.25)
where $l_{yx} < 0$ and $l_{xy} > 0$ by assumption. Although $q'_x$ is now smaller than it was before the presumed inflow of foreign capital and the resultant interregional factor flows, i.e. $q'_x < q_x = n$, while $q'_y > q_y = n$, the growth rate of output per worker is again zero in both regions.

Finally, it is of interest to note that the corresponding growth rate of the total or combined regional output is higher now than it was before the foreign capital inflow into region Y; and similarly, since the natural growth rate of the total labour force is exogenously given, it follows that the growth rate of output per worker in the economy as a whole will also be higher now than before. The reason for this is, of course, similar to that pertaining to the comparative static analysis of the previous section: regional factor mobility makes it possible for a perfectly competitive economy to utilize its growing resources most efficiently.

6. **CONCLUSION: The Critical Assumptions**

There is nothing inherent in the neoclassical theory of general equilibrium that provides an explanation of the emergence, let alone continued existence, of the problem of economic underdevelopment. In order to analyze the development problem from a purely neoclassical perspective, it becomes necessary to consider various exogenous changes which have the effect of

29. These propositions are further explored in Appendix 3 below.
distorting, at least temporarily, the otherwise permanent configuration of competitive prices and quantities. But such disarrangements will, of course, be automatically eliminated by the equilibrating mechanisms operating within a perfectly competitive economy. Indeed, if one were to accept without question the neoclassical theory in its pure and unadulterated form, there would be no such field as development economics.

It is generally known, of course, that the predictions of the neoclassical model have not come true. There is enough evidence indicating that the development question has become progressively more problematical over the years. Capital has not flowed to the relatively backward, predominantly agricultural sectors to any significant extent, and although many labourers have migrated to the rich, industrial sectors, only a relatively small number has managed to obtain adequate employment. Notwithstanding a large and growing supply of labour, entrepreneurs in the industrial sectors have tended to use relatively capital-intensive production techniques. The main reason for this pattern of resource allocation lies in the generally imperfect nature of the market, especially in developing countries: constant returns to scale are the exception rather than the rule; production functions are characterized by a limited degree of factor substitutability; capital and labour are neither completely divisible nor perfectly mobile over space; competitive markets are conspicuous by their absence, while price and quantity adjustments are often disequilibrating due to insufficient knowledge on the part of producers and consumers alike. It is largely for these reasons that the arch proponents of the

30. For a devastating critique of the neoclassical theory, see Richardson (1973, ch.1).
(pure) neoclassical theory may be said to be "guided much more by logical curiosity than by a taste for relevance".  

CHAPTER 2

ECONOMIES OF SCALE, TECHNOLOGICAL PROGRESS AND THE FACTOR-PROPORTIONS PROBLEM

In this chapter we shall consider several theories of economic underdevelopment which focus attention on production conditions in developing countries. What these theories have in common is an analytical approach which essentially amounts to relaxing one or more of the basic assumptions underlying the production function of the neoclassical two-sector model. In section 1 below it is argued, for example, that the theory of cumulative causation effectively introduces increasing returns to scale in at least one sector or region of the economy; and as may be expected, its predictions turn out to be vastly different from those associated with the neoclassical assumption of uniform constant returns. Similarly, section 2 attempts to compare the dynamic characteristics of increasing returns with the convergency properties of the neoclassical theory of growth. Section 3 introduces the fixed-proportions production function, which forms the basis of Richard Eckaus's (1955) theory of the "factor-proportions problem". In section 4 we consider the implications for the factor-proportions problem of Harvey Leibenstein's (1960) suggestion that the impact of technological progress is generally limited to those sectors of the economy using relatively capital-intensive production techniques. Finally, sections 5 and 6 examine the general policy implications of the existence of fixed factor proportions, both from a static and dynamic perspective.
1. **Differential Returns to Scale: The Theory of Cumulative Causation**

The name most closely associated with the theory of cumulative causation is that of Gunnar Myrdal.\(^1\) In some of his major works\(^2\) he has consistently rejected the (neoclassical) "notion" of a stable equilibrium: "I feel, indeed, very much in line with ordinary common sense when I stress that in the normal case circular causation is a more adequate hypothesis than stable equilibrium for the theoretical analysis of a social progress".\(^3\) According to Myrdal, most social progresses are inherently unstable insofar as the behavioural relationships between their constituent variables are "cumulative because of circular causation". This hypothesis is said to be especially relevant to the process of economic development which "... normally tends to increase, rather than to decrease, the inequalities between regions".\(^4\)

Any given increase in a region's income is assumed to have a dual effect on other regions, in the form of the well-known spread and backwash effects. The claim that the latter effect usually outweighs the former is due largely to the existence of regional differences in economies of scale.\(^5\) Specifically, suppose an autonomous increase in the income of, say, region Y induces increasing returns and raises its marginal productivities and

---

1. But see also Hirschman (1957; 1958).
5. See Myrdal (1957, p. 27), Kaldor (1970) and Dixson and Thirwall (1975).
factor prices relative to those in region X. If factor mobility is assumed to depend on the corresponding price differentials only, then production factors will move in one direction only setting in motion a cumulative process of income expansion in region Y at the expense of region X. In the absence of "contervailing forces", this process could continue ad infinitum augmenting regional differences in scale economies, real income and employment.

Such cumulative processes have, of course, been recognized long before and made implicit in the familiar notion of the "viscous circle"; similarly, growing interregional inequalities have been explained in terms of the "cumulative" nature of the Keynesian multiplier process. Nevertheless, when it was first applied to the problem of economic underdevelopment, the theory of cumulative causation initiated a fairly radical departure from the ruling neoclassical orthodoxy. The basic idea behind the theory has recently re-emerged in various writings associated with the so-called dependency or "neo-Marxist" school. These writers contend that past colonial empires and such "neo-colonialists" as the multinational corporation have both played an important role in initiating and perpetuating a process of "development of underdevelopment". Through various forms of international trade, investment and technological transfer, for example, multinational companies establish a so-called "enclave" economy within the

6. See, for example, Nurske (1953, p. 4).
8. Myrdal (1944)
typical developing country, which eventually becomes fully integrated with the "international capitalist system". The chief function of the enclave is that of profit (or "surplus value") maximization, which it does inter alia by partially destroying traditional (handicraft) industries and retarding or distorting indigenous processes of social and economic change. Implicit in the dependency argument is the belief that backwash or polarization effects far outweigh what spread effects may emanate from the enclave economy, at least during the early stages of development.

Whatever the reasons may be, it is difficult indeed to reject the hypothesis of a "long swing in the inequality characterizing the secular income structure"; that is, much of the available evidence suggests that regional income inequalities are greatest in the poorest countries of the world. By the same token, however, it must be conceded that beyond a certain threshold level of development, such regional inequalities tend to diminish as the spread effects gradually gain in importance over the backwash efforts. This is presumably what Hirschman (1958, p. 189) had in mind when he wrote: "Myrdal's analysis strikes me as excessively dismal ... he fails to recognise that the emergence of growing points and therefore of differences in development between regions and between nations is inevitable and is a condition of further growth anywhere".

13. See also Bauer (1963) and Bauer and Yamey (1967).
It seems worth analyzing the general equilibrium implications of cumulative causation in terms of the two-sector model developed in the previous chapter. For this purpose, it is useful first to note that Myrdal (1957, p. 27) himself referred to the phenomenon of economies of scale "in the broadest possible sense" — that is, it includes both internal economies and such external economies as ownership, technical and public goods externalities. The existence of internal or external economies implies, of course, that the average cost of production is falling and exceeds the marginal cost over the relevant range. Such a situation either leads to complete market failure under perfectly competitive conditions, or it gives rise to monopolistic or oligopolistic pricing behaviour according to which price is set above the marginal cost of production; in either case, "the correspondence between market-directed and welfare-maximising allocation fails".

In what follows here, however, we shall for expositional purposes ignore the general equilibrium properties of imperfectly competitive markets, and assume instead that price is set equal to the marginal cost of production — for example, by a Lange-Lerner type of bureaucracy. While the role of monopoly in economic development will be more thoroughly examined in Chapter 3, it is worth noting here that the presumed existence of (competitive)

14. See, for example, Bator (1958) and Koutsoyiannis (1979, ch. 23).
15. Ibid.
17. See Bator (1957).
shadow prices is unlikely to affect our subsequent analysis of increasing returns to any significant extent: "The marginal-rate-of-substitution conditions ... retain their validity, and the solution still gives out a set of shadow prices, or decentralized responses, which result in the (optimal) configuration of inputs, outputs and commodity distribution". 18

Starting from an initial equilibrium situation in Figure 2.1(a), let us introduce increasing returns into region Y only, such that they occur at a uniform rate over the full range of inputs and outputs; specifically, suppose that equal proportionate increases in capital and labour inputs give rise to constant, but proportionately larger, increases in output over the relevant range. 19 This implies that the distance between consecutive Y-isoquants become progressively smaller as one moves farther away from Y's origin; or put differently, the effect of increasing returns is to shift the Y-isoquants successively closer to O_y, for example from Y_2 and Y_4 to Y'_2 and Y'_4 respectively. This also explains why the transformation curve in Figure 2.1(b) becomes generally less concave than before, and why it may even change its shape altogether from being concave to being convex to the origin; indeed, it is quite possible that the degree of convexity of the transformation curve may be such as to yield a so-called corner solution. 20

18. Ibid. p. 408.
19. See also Koutsoyiannis's (1979, pp. 79-81) treatment of increasing returns occurring at constant and variable rates.
FIGURE 2.1
These possibilities are illustrated in Figure 2.1(b). Firstly, if increasing returns were to occur at a relatively modest rate in region Y, then the transformation curve might shift from TT' to its less concave counterpart ST', in which case the equilibrium output mix would change from point E to point F. Alternatively, should increasing returns occur at a sufficiently high rate, then the transformation curve could become convex to the origin, such as RT' in Figure 2.1(b). In this case a corner solution may be established at a new equilibrium output mix, for example point R where the community indifference curve labelled W₃ intersects RT'. The significance of the latter equilibrium is simply the fact that the total supplies of capital and labour are now being used in the production of Y only: assuming "appropriate" demand conditions, the increase in the marginal productivities (and factor prices) has been such as to induce X's entire factor supplies to migrate to region Y, thus raising output in Y considerably while lowering it to zero in region X.

2. Differential Returns and Dynamic Equilibrium

The dynamic properties of differential returns to scale may be explained by means of a simple extension of the neoclassical growth model developed in the previous chapter.21 First, recall the neoclassical growth equations,

21. See also Eltis (1973, ch. 11).
where the underlying production functions are assumed to portray constant returns to scale. If we now introduce increasing returns into region $Y$, then

$$y' = \lambda \{ \beta k_y + (1 - \beta) l_y \}$$

where $\lambda > 1$ and indicates the degree of increasing returns. Now, it can be shown that steady state growth implies that $y' = k_y$;\(^{22}\) or

$$y' = \frac{\lambda(1 - \beta) l_y}{(1 - \lambda \beta)}$$

where $\lambda (1 - \beta)/(1 - \lambda \beta) > \lambda$. Increasing returns thus have the effect of raising the neoclassical growth rate by a multiplier of $\lambda (1 - \beta)/(1 - \lambda \beta)$; that is, it "...multiplies the growth that is due to any other source, such as growth in the industrial labour force..... thus magnifying the advantages countries obtain from other sources of growth".\(^{23}\)

It is perhaps worth noting here that the latter conclusion is consistent with Verdoorn's (1947) law, which has been frequently verified empirically,\(^{24}\) and which simply states that there is a strong correlation between the rate of growth of labour productivity and the rate of output growth; for example,

---


23. Eltis (1973, p. 258). It might be added that if increasing returns were to occur at an increasing rather than constant rate, then this would provide a further magnifying effect on the growth rate of $Y$.\(^{24}\)

24. See, for example, Kaldor (1966).
\[ y - l_y = \pi y \]  \hspace{1cm} (2.5)

or \[ y = \frac{l_y}{1 - \pi} \] \hspace{1cm} (2.5')

where \( \pi \) is known as the "Verdoorn coefficient". If \( 0 < \pi < 1 \), as seems likely, then the (neoclassical) growth rate is again raised by a multiple, in this case of \( 1/(1 - \pi) \).

Finally, recall that in our present example regions X and Y are assumed to experience constant and increasing returns to scale, respectively. The steady state is accordingly characterised by \((x - l_X) = 0\) and \((y - l_Y) > 0\). This implies, generally, that the growth rate of output per worker will be positive in those regions experiencing increasing returns to scale, while it will be zero (or negative) for regions subject to constant (or decreasing) returns to scale. Consequently, the predictions of the differential returns hypothesis turn out to be quite different from the convergency properties of the neoclassical growth theory: Myrdal's theory of cumulative causation envisages a process of regional divergence in the output level per worker.

3. The "Factor-proportions Problem"

Although the "factor-proportions problem" was introduced to the literature on economic development by Richard Eckaus (1955), it has in fact a long, time-honoured history beginning with such classical writers as Ricardo (1951, pp. 76-77) and Mill (1965, pp. 63-65). Subsequently, Pareto (1887, p. 717) described it

25. See also Dorfman (1953) and Fukuoka (1955).
as follows: "..... in order to produce a given amount of silks, one requires an area of land to erect a factory, but, afterwards, even if one doubles this area, without increasing the other capital goods, the product will not be increased at all". Limited factor substitutability has since been viewed as an important source of unemployment during the crisis years of the 1930's\textsuperscript{26} — notwithstanding the emphasis given at the time to the Keynesian notion of deficient demand. It is interesting to note too that Hicks (1947) considered the existence of fixed factor proportions to be a critical determinant of the twin problems of unemployment and inflation that existed in Europe during the immediate postwar period.

In the field of development economics, it has long been felt that the problems of factor underutilization generally, and of labour unemployment in particular, are at least partly caused by the fact that entrepreneurs are faced with a limited range of relatively capital-intensive production techniques. The origin of this type of production function lies with what Hans Singer (1970-71, p. 64) has called the system of international technological dualism —".... the fact that knowledge is accumulated by the richer countries, in the richer countries and in respect to the problems of the richer countries"\textsuperscript{27}. Such a monopoly of knowledge enables the developed countries to control both the volume and composition of technical innovations.

\begin{thebibliography}{9}
\bibitem{26} Keynes (1936, p. 296), Kaldor (1938, pp. 643-644) and Robinson (1965, pp. 78-80).
\bibitem{27} See also Singer (1975), Streeten (1971), Stewart (1972; 1974), Cooper (1972) and Helleiner (1975).
\end{thebibliography}
and inventions in the world at large, as well as their eventual transfer to developing countries through international trade and investment. However, since most innovations are initially planned and designed in developed countries where labour tends to be relatively scarce and expensive, their application generally requires large quantities of capital relative to labour.\textsuperscript{28} It is the application (without adaptation) of such innovations in developing countries which has led Eckaus (1955), Higgins (1958, ch. 14) and others\textsuperscript{29} to observe that despite the relative abundance and cheapness of labour, production in these countries is often characterized by fixed factor proportions of a capital-intensive nature.

Both Eckaus and Higgins divide the typical developing economy into two sectors or regions of economic activity, namely a capitalist or industrial sector and a traditional, predominantly agricultural sector. In the former sector there is either in fact, or entrepreneurs perceive there to be, fixed technical coefficients of production: "If managers and technicians, used to particular methods of production in Western countries which they accept without question as superior, do not look for alternative techniques more suitable to the factor endowment, the effect is the same as if coefficients were technologically fixed."\textsuperscript{30}

\textsuperscript{28} Streeten (1971) and Arrighi (1970) maintain that the choice of techniques may be influenced by a desire on the part of multinational companies to economise on the use of "foreign" labour, whether it be skilled or unskilled.

\textsuperscript{29} See Seers (1972) and the references listed in footnote 27.

\textsuperscript{30} Higgins (1968, p. 301 (italics added)). See also Eckaus (1955, p. 353).
Reverting to our two-sector model, let us suppose that the production of \( Y \) is in fact characterized by fixed factor proportions;\(^3\) that is,

\[
Y = \min (\sigma K^a_y, vL^a_y)
\]

(2.6)

where \( K^a_y \) and \( L^a_y \) are the available supplies of capital and labour in the capitalist sector, \( Y \), and \( \sigma \) and \( v \) are fixed (or "selected")\(^2\) technical coefficients representing the output/capital and output/labour ratios respectively. If \( \sigma K^a_y < v L^a_y \), then

\[
Y = \sigma K^a_y = v(L^a_y - U_y)
\]

(2.6′)

where \( U_y \) represents labour unemployment. Let \( K^a_y = K_y \) and \( (L^a_y - U_y) = L_y \) be the quantities of capital and labour actually used in the production of \( Y \) respectively; then

\[
Y = \sigma K_y = vL_y
\]

(2.6″)

or

\[
K_y = (v/\sigma) L_y
\]

(2.7)

represents the fixed-proportions production function for sector \( Y \). This is shown in Figure 2.2 by the ray \( O'yA' \) whose slope, \( v/\sigma \), measures the (fixed) capital/labour ratio used in the production of \( Y \).

Similarly, production techniques in the traditional sector, \( X \), are assumed to be variable albeit within a limited range of factor proportions only; that is, sector \( X \)'s factor absorption

---

\(^3\) The alternative case in which factor proportions are variable but nevertheless perceived to be fixed, will be discussed in section 5.

\(^2\) They are referred to as selected coefficients if there are more than one production process available. (see Allen (1967, p. 36)).
Figure 2.2

\[ K_y = \frac{v}{\sigma} L_y \]
capacity is constrained by a limited degree of factor substitutability. Such a production function is shown in Figure 2.3, where the ridge lines, $O'H'$ and $O'G'$, represent loci of input proportions for which the respective marginal productivities of capital and labour are zero; for example, if labour is added to the fixed quantity of $O'K'$ capital units, then its marginal product will soon fall to zero at point $R'$, beyond which additional labour units will simply remain unemployed or underemployed.  

Let us now combine Figures 2.2 and 2.3 to form the Edgeworth box diagram shown in Figure 2.4(a) below. Ignoring for the moment the ray $O'J'$, the efficiency locus or path of maximum total output may be derived simply by finding the maximum output of $X$, for every given $Y$-output along the ray, $O'A'$: for example, if the output of sector $Y$ is at level $Y_3$, then the maximum attainable $X$-output is given by $X_2$ along the ridge line, $O'H'$. Applying this procedure to each level of output $Y$, we obtain the efficiency locus, consisting of the line segments $O'E'$ cum $A'E'$ and $E'O_y$ in Figure 2.4(a). This information may now be used to trace out the corresponding transformation curve, $AEQ_y$ in Figure 2.4(b), where the segments $AE$ and $EO_y$ accord with the segments $O'E'$ cum $A'E'$ and $E'O_y$ of the efficiency locus, respectively. It should be noted that while output combinations along $EO_y$ indicate full employment of both factors in the economy as a whole, the output combinations given by $AE$ represent at least some labour unemployment.

---

33. While the distinction between unemployment and underemployment will be made more explicit in Chapter 3, suffice it to mention here that underemployment generally refers to a situation in which a limited amount of work is being shared out between an excessively large number of available workers.

34. This diagram is similar to the one used by Eckaus (1955, p. 369) in his analysis of the factor-proportions problem.
FIGURE 2.3
FIGURE 2.4
In purely theoretical terms, the wage rate under competitive labour market conditions will be zero for output combinations lying on the labour-unemployment segment, AE in Figure 2.4(b), corresponding to the input combinations given by O'E' cum A'E' in Figure 2.4(a). However, it is surely more realistic to assume that variations in the wage rate are constrained institutionally by a given minimum (and positive) value. The implications of such a minimum wage for the factor-proportions problem may be explained with the aid of Figure 2.4. Let the minimum wage be given by the slope of the factor price line aa' passing through point M' on sector Y's expansion path, O' A'. Assuming profit maximizing behaviour in both sectors, the constrained or "limited" factor input ratio for sector X is then determined at point N'; while the slope of price line bb' will be equal to that of aa'. Similarly, the assumption of constant returns to scale ensures that the ray O'N', extended to J', represents a locus defined by factor price lines whose slopes all equal that of bb' or aa': in other words, the imposition of the minimum wage reduces sector X's feasible output region from O'G'O'H' to O'G'O'H' in Figure 2.4(a); and similarly, it causes the efficiency locus to shift

35. While this possibility is also briefly considered by Eckaus (1955, pp. 373-374), the role of factor price distortions in economic development will be more thoroughly explored in Chapters 3 and 4.


37. See Baldry (1980, pp. 139-141).
from $O_x'E_x' cum A'E_x'$ and $E_x'O_y'$, to $O_x'F_x' cum A'F_x'$ and $F_x'O_y'$, and the
transformation curve from $AEO_y$ to $AFO_y$. Accordingly, the
effect of the minimum wage is to lengthen the labour-unemployment
segment of the transformation curve from $AE$ to $AF$ in Figure 2.4(b).

Let us now examine the role played by demand conditions in
determining the magnitude of the factor-proportions problem. Consider,
for example, the community indifference curves labelled
$W_1$ and $W_2$ in Figure 2.5 below, indicating a relatively strong
community preference for commodity $Y$. Equilibrium is established
at $Z$ where the community indifference curve, $W_2$, is tangent to
the labour-unemployment segment $AE$ of the constrained transformation
curve. But since the latter equilibrium implies the existence of
labour unemployment, it follows that the community is faced here
with a conflict between maximum social welfare on the one hand,
and overall full employment on the other: although it is technically
possible to attain full employment of both factors at the output
mix $E$, this would clearly imply a decrease in the level of social
welfare from $W_2$ to $W_1$ in Figure 2.5. Moreover, it is Eckaus's
contention that the transformation curve of the typical developing
country consists mainly of labour-unemployment segments such as $AE$
in Figure 2.5; and if this is indeed the case, then the existence
of a trade-off relationship between full employment and maximum
welfare becomes a very real possibility: "..... technology, factor
endowments and final demand may combine in ways which make it very
difficult for underdeveloped areas to solve their problems of
unemployment and underemployment".

38. It should be added that the minimum wage is only effective or
binding for those input and output ratios lying on the labour-unemploy-
ment segment $O_x'F_x' cum A'F_x'$ in Figure 2.4(a), or $AF$ in Figure 2.4(b). It
thus follows that the wage rate will be higher in both sectors for those
ratios lying on the remaining segments of the efficiency locus and
transformation curve.
4. Inappropriate Technology and the Factor-proportions Problem

Harvey Leibenstein (1960) has put forward a theory of technological progress which has an important bearing on the factor-proportions problem. His basic hypothesis is that technological progress occurs mainly in the economically advanced sectors or regions, where it gives rise to relatively small and gradual improvements in the quality of given types of capital equipment. In the traditional sector where the capital/labour ratio and the wage rate are low, however, the likelihood of recognizing opportunities for such marginal innovations is limited indeed, while the scale of operations may also be too small to support even small technical improvements in the capital stock. 41 Accordingly, technological progress (of a capital-using type) is more likely to raise factor productivities in those sectors of the economy using relatively capital-intensive production techniques. 42

Leibenstein implicitly assumes the existence of a single, uniform production function for the economy as a whole. Consider, for example, the linear production function consisting of a limited number of fixed-proportions processes shown in Figure 2.6. Technological progress is assumed to raise the productivity of those production processes for which the capital/labour ratio is relatively high rather than low; that is, it causes a shift in the isoquant from \( Q_1 \) to \( Q_1' \). 43 But since the latter processes are used in one or only a few sectors of the economy, Leibenstein

---

41. This is precisely the reasoning underlying Todaro's (1969) view that production functions exhibit greater returns to scale for capital-intensive as opposed to labour-intensive production techniques.

42. See also Bruton (1965, p. 41) and the references listed in footnote 27 above.

43. See Leibenstein (1960, p. 354).
FIGURE 2.6
maintains that while technological progress has the greatest effect on the choice of techniques in the more capital-intensive sectors, it leaves production techniques in the more labour-intensive sectors virtually unaffected.

When applied to the two-sector model above, Leibenstein's proposition essentially amounts to the introduction of capital-using technological progress in the capital-intensive sector of the economy. This may be illustrated with the aid of Figure 2.7. The initial efficiency locus is again given by the line segments $0_x'A_x'CE'$ and $E_0'O_y'$ in Figure 2.7(a), from which is derived the transformation curve $AEO_y$ in Figure 2.7(b). Demand conditions initially determine a full-employment equilibrium at $E$, where the community indifference curve labelled $W_1$ is tangent to the transformation curve. Now suppose a given technological innovation enables sector $Y$ to employ fewer units of both capital and labour in the production of any given level of output $Y$; and similarly, the fact that the innovation is assumed to be capital-biased means that sector $Y$ now uses more capital relative to labour than before. The $Y$-isoquants accordingly shift from $Y_1'$ and $Y_2'$ to, say, $Y_1'$ and $Y_2'$ in Figure 2.7(a). The corresponding shift in the efficiency locus results from the fact that production process $O_y'B_y'$ is assumed to "dominate" process $O_y'A_y'$: for example, if sector $X$ produces the output level, $X_2'$, then the maximum attainable $Y$-output is given by $Y_3'$ along ray $O_y'B_y'$, rather than by $Y_2'$ along ray $O_y'A_y'$. Repeating this process yields a new efficiency locus, comprising the line segments $0_x'C_x'B_y'$ and $B_y'F_y'O_y'$. This implies, in turn, a shift in the transformation curve from $AEO_y$ to $BFO_y$ in Figure 2.7(b), thus effectively lengthening the labour-unemployment segment from $AE$ to $BF$. Likewise, the
equilibrium output mix moves from \( E \) to \( Z \), where the community indifference curve, \( W_2 \), is tangent to the new transformation curve. However, since the latter equilibrium is tangent to the labour-unemployment segment, \( BF \), it follows that the technological innovation has in fact given rise to labour unemployment in the economy as a whole.

Generally, the above example serves to illustrate how an inappropriate technological innovation — or one that is ill-suited to the factor proportions in question — could aggravate the employment problem within a typical developing country. But the unemployment thus created is not, of course, the result of technological progress \emph{per se}: rather, it arises from the fact that the innovation is being applied to production processes characterized by limited factor substitutability.

5. \textbf{Policy Implications of the Factor-proportions Problem}

The solution to the factor-proportions problem depends in part on whether such proportions are in fact technologically fixed, or whether they are actually variable, but nonetheless perceived to be fixed. If the latter illusion applies, for example, then there is clearly a need to inform entrepreneurs of the availability of alternative production techniques, at least one of which is likely to improve the profitability of their business. This proposition is illustrated diagrammatically in Figure 2.8. The (representative) firm's actual production function is given by solidly drawn isoquants
FIGURE 2.8
labelled $Y_1$ and $Y_2$. For reasons discussed earlier, however, the firm believes that the ray, $O'Y'$, represents its only feasible production process. Accordingly, if price line $aa'$ indicates the firm's cost constraint, then equilibrium will ostensibly occur at point $E'$, where $aa'$ is tangent to the (perceived) L-shaped isoquant labelled $Y_1$. But $E'$ does not, of course, represent the true profit maximizing equilibrium, insofar as it is indeed technically feasible to produce the higher output level $Y_2$, by means of the input combination given by $E$. Likewise, if all firms could be persuaded to use such optimal factor proportions indicated by point $E$ in Figure 2.8, then it should be possible for the economy as a whole to produce along its (true or unconstrained) transformation curve, and hence to achieve full employment of both production factors at all times.

Alternatively, the existence of fixed factor proportions calls for an entirely different set of policy measures. In particular, it would be necessary here to devise a feasible range of new production techniques more suited to the factor endowment and demand conditions of the country in question. The introduction of a new, more labour-intensive production process would, for example, shorten the labour-unemployment segments of the efficiency locus and the transformation curve, and in so doing increase the likelihood of the economy attaining a full employment (static) equilibrium. This is illustrated in Figure 2.9 where, as before, the initial efficiency locus is given by the line segments

44. See, for example, Pack and Todaro (1969), Sutcliffe (1971, ch.5), Morawetz (1974), Timmer et al. (1975) and Pack (1976).
FIGURE 2.9


(b) Diagram with axes O_y, X, Y, C, A, M, Z, W_2, N, U, E, W_1, P, O_x.
O'E' cum A'E' and E'O' in Figure 2.9(a), from which is derived the transformation curve AEQ_y in Figure 2.9(b). Let us now introduce a new production process in sector Y, represented by the ray O_y'C' in Figure 2.9(a). This implies, of course, that sector Y is now able to use either process O_y'A' or process O_y'C', or indeed some combination of the two processes. Since the derivation of the new efficiency locus and transformation curve is similar to the derivation procedure used previously, suffice it to mention here that the former now consists of the segments O_x'M' cum C'M', M'N', N'P' and P'O_y'; and likewise, the new transformation curve consists of the similarly lettered segments, CM, MN, NP and PO_y.

It is evident that the labour-unemployment segment has been considerably shortened, for example from AE to CM in Figure 2.9(b). Likewise, the fact that the new equilibrium output mix now occurs at Z, implies that the economy has moved from a position of labour unemployment at U, to one representing both full employment and maximum social welfare. Consequently, the introduction of the new, more labour-intensive production technique has effectively removed the trade-off relationship that existed between full employment and maximum welfare.


We finally turn to the dynamic policy implications of the factor-proportions problem. Consider, for example, the derivative
of the production function for \( Y \) (i.e. equation \( 2.6'' \) above):

\[
dY = \sigma dK_y = v dL_y
\]  

(2.8)

Substituting \( \sigma = Y/K_y \) and \( v = Y/L_y \) into (2.8), we get:

\[
dY = Yk_y = Yl_y
\]  

(2.8')

or

\[
y = k_y = l_y
\]  

(2.9)

where \( y, k_y \) and \( l_y \) represent the growth rates of output, capital and labour employment, respectively. Suppose now \( l_y < l^a_y \), where \( l^a_y \) indicates the growth rate of the labour supply in sector/region \( Y \), assumed to be exogenously determined. The corresponding growth rate of labour unemployment is then given by \( u_y = l^a_y - l_y \).

In contrast to the neoclassical theory of growth, it is highly unlikely that the present economy will attain full employment automatically. This follows immediately from the fact that:

\[
dK_y = s_y Y
\]  

(2.10)

\[
or \quad k_y = s_y
\]  

(2.11)

where \( s_y \) is the marginal savings propensity in region \( Y \). Accordingly, in the absence of interregional factor movements, it would just not be possible to raise \( k_y \) (and \( l_y \)) in view of the presumed constancy of \( \sigma \) and \( s_y \). Should production factors be mobile, however, it would be possible, at least, to attain overall full employment if labour were to migrate from region \( Y \) to region \( X \) and/or if capital flowed in the opposite direction.
But this would, in turn, require that \( w_x > w_y \), \( r_x < r_y \) and \( k_x > l_x \). Such an equilibrium would be purely coincidental, since the existence of fixed factor proportions implies that the various growth rates and factor prices are determined independently from one another.\(^{45}\) An equilibrium thus established would be tantamount to what Joan Robinson (1965) has called a "golden age", or "a mythical state of affairs not likely to obtain in any actual economy".

The dynamic policy implications of fixed factor proportions are essentially threefold. In the first place, a strong case could be made for the implementation of labour-using technological innovations and inventions, the effect of which would be similar to that associated with the comparative static analysis of the previous section: for example if \( l_y^a > l_y \), then the effect of such a policy should be to raise the growth rate of labour employment by \( u_y \), whence it follows that \( k_y = l_y^f = l_y^a \), where \( l_y^f = l_y + u_y \).

Secondly, it may be possible to reduce the rate of population growth, and hence the growth rate of the labour supply too, if it is feasible to lower the average fertility rate in the country.\(^{46}\) This could in principle be done in several ways. One possibility would be to use the existing communications and educational media as a means of informing the general public about the economic advantages of small-sized families. Another is the establishment of family planning clinics whose task it would be to give advice on available birth-control measures, while at the same time offering abortion services and a suitable range of contraceptives. Likewise, many

---

45. See, for example, Singh (1966, ch. 2).
46. See, for example, King (1974) and Todaro (1977, pp. 157-161). It has been estimated (Enke (1967) and Zaidan (1971)) that the social returns of such a policy are likely to be much higher than those associated with most other forms of investment.
countries have recently introduced various financial incentives and disincentives in an attempt to encourage small-sized families, birth spacing and delayed marriages. In terms of our earlier notation, the desired effect of these measures would be to lower the growth rate of the labour supply by \( u_y \), so that \( k_y = \ell_y = \ell^a_y \) where \( \ell^a_y = (\ell^a_y - u_y) \).

A third remedy to the factor-proportions problem consists in raising the growth rate of the capital stock; given fixed factor proportions, for example, this should cause a corresponding increase in the growth rate of labour employment. Increased capital growth could be achieved through (i) the creation of more and better savings institutions and the implementation of appropriate monetary policy measures; (ii) the use of various direct and indirect taxes; and (iii) the encouragement of foreign investment. Here again the desired effect should be to raise the growth rates of capital and labour employment by \( u_Y \), whence \( k'_y = \ell'_y = \ell^a_y \), where \( K'_y = k'_y + u_y \) and \( \ell'_y = \ell'_y + u_y \).

---

47. See Todaro (1977, pp. 158-159).
48. For a dissenting view, see Bauer (1981).
49. See Higgins (1968, ch. 23) and Meier (1976, pp. 267-270).
50. See Thirlwall (1972, ch. 8; 1976, ch. 2).
CHAPTER 3

MARKET IMPERFECTIONS AND ECONOMIC UNDERDEVELOPMENT:
THE CLOSED ECONOMY

Much of the postwar literature on economic development has focused on the imperfectly competitive structure of markets in general. Singer (1950) and Prebisch (1959) for example, have shown how differences in the structure of markets between the developed and developing countries could turn the terms of trade against the latter. Myint (1954), Eckaus (1955) and others\(^1\) have again pointed to numerous barriers to entry which limit the degree of competition in both the product and the factor markets in developing countries. Similarly, Lewis (1954), Ranis (1973), Todaro (1971; 1977) and several recent studies by the International Labor Office (ILO) (1970; 1972), have observed that domestic factor markets are characterized by the existence of minimum wage legislation and also by distorted capital prices. The effect of such price distortions generally is to limit the growth of employment and income opportunities, which is often aggravated by certain patterns of spatial factor mobility: Myrdal (1957) and Todaro (1969) both argue, for example, that the mobility of production factors tends to be spatially disequilibrating insofar as it aggravates urban unemployment and widens per capita income differentials within developing countries. It is indeed this very spectrum of "market imperfections" which prompted Todaro (1971, p. 396) to remark: "... it is now becoming

\(^1\) See for example, Johnson (1962) and Little et al. (1970).
painfully apparent how the conventional wisdom of economic theory, which placed top priority on the rapid accumulation of capital as the key to successful economic progress in the 1950's and early 1960's, has led to the serious employment predicament of the 1970's).

In this chapter we shall consider in some detail the employment and the general equilibrium implications of market imperfections — that is, of the various price distortions that pervade the product and the factor markets of a "typical" developing country. Product market imperfections mainly refer to the existence of various artificial barriers that restrict entry by new firms into relatively profitable industries, thus limiting output and raising prices above their respective competitive levels. But such barriers do not give rise to the unemployment problem as such — they merely "aggravate the factor disequilibrium" that already exists in the factor markets.  

In what follows in this chapter, we shall be chiefly concerned with the employment effects of factor market imperfections; only section 1 below refers to the role played by monopolies in the product markets of developing countries. Sections 2 and 3 consider Lewis's dualist theory and Michael Todaro's model of rural-urban migration, respectively, and specifically try to place them within the general context of factor market imperfections. Section 4 examines the relatively recent phenomenon of the urban informal sector, emphasizing the fact that its emergence and continued survival depend in part on the existence of various price distortions in the urban areas of

developing countries. Finally, section 5 discusses the general policy implications of market imperfections.

1. Monopoly, Monopsony and Economic Underdevelopment

The presence of foreign-owned monopolies and/or oligopolies has been a characteristic feature of the "opening-up" process in developing countries. These firms have generally been able to determine both the price and the quantity supplied of their product or service, whether it be for export or for intermediate or final demand purposes in the domestic market. It is sometimes argued that monopoly is a necessary condition for economic development during its initial stages, in view of the small size of the domestic market, the relative costliness of operations and the need to initiate and implement technological inventions and innovations. On this view, the general role of monopoly in economic development may be likened to that of a natural monopolist; that is, the entire market can be supplied by one firm only, operating at comparatively low unit cost of production.

One of the problems with most monopolies is that they tend to exist long after the need for them, as set out above, has disappeared. Such firms often operate in markets large enough to be served by a great many small-sized, competitive firms, with the result that the market output is bound to be smaller.

3. See, for example, Myint (1954), Barnett and Muller (1974) and the references on "dependency" given in footnote 9 of Chapter 2.

4. Tugendhat (1973, Ch.6), Lall (1973; 1975), Hansen (1975) and Vaitos (1976).

and price higher than they would have been under conditions of perfect (or free) competition. Moreover, these 'structural' imperfections are not limited to product markets only: "... in a typical process of 'development', the backward peoples have to contend with three types of monopolistic forces: in their role as unskilled labour they have to face the big foreign mining and plantation concerns who are monopolistic buyers of their labour; in their role as peasant producers they have to face a small group of exporting and processing firms who are monopolistic buyers of their crop; and in their role as consumers of imported commodities they have to face the same group of firms who are monopolistic sellers of distributors of these commodities".6

Reverting to the two-sector analysis of our earlier chapters, let us suppose that sector $Y$ represents a monopolist while $X$ is one of a (large) number of competitive firms producing a similarly lettered commodity. The profit function of $Y$ is given by:

$$\Pi_y = Q_y P_y - C_y$$  \hspace{1cm} (3.1)

where $C_y$ represents the total production cost of $Y$. Since $P_y = P_y(Q_y)$ by assumption, and $C = C(Q_y)$, we may write (3.1) as follows:

$$\Pi_y = Q_y P_y(Q_y) - C(Q_y)$$  \hspace{1cm} (3.1')

The first-order condition for profit maximization requires that the first derivative of (3.1') equals zero; that is,

$$d\Pi_y/dQ_y = P_y(Q_y) + Q_y(dP_y/dQ_y) - dC_y/dQ_y = 0$$  \hspace{1cm} (3.2)

or

$$P_y(Q_y) + Q_y(dP_y/dQ_y) = dC_y/dQ_y$$  \hspace{1cm} (3.2')

indicating the familiar equality between marginal revenue and marginal cost. The fact that \( \frac{dP_y}{dQ_y} < 0 \) means, of course, that \( P_y \) exceeds the marginal cost of production.

Now, the general equilibrium implications of monopoly follow from the (given) equality between the marginal rate of product transformation on the one hand, and the ratio of the marginal cost of \( X \) to the marginal cost of \( Y \) on the other; or

\[
\frac{dQ_y}{dQ_x} = \frac{dC_x/dQ_x}{dC_y/dQ_y} \tag{3.3}
\]

where \( \frac{dC_x/dQ_x} \) represents the marginal cost of \( X \).\(^7\) Accordingly, since \( P_y > \frac{dC_y/dQ_y} \) while \( P_x = \frac{dC_x/dQ_x} \) by assumption, we have

\[
\frac{dQ_y}{dQ_x} > \frac{P_x}{P_y} \tag{3.4}
\]

which indicates simply that the third or 'top level' condition for Pareto optimality is not being complied with. This is illustrated in Figure 3.1 by the difference between the slope of the transformation curve, \( TT' \), and the commodity price line, \( P_1P_1' \), passing through point \( F \): specifically, in contrast to

\(^7\) This is easily proved by means of the following familiar identities:

\[
\frac{\partial Q_x}{\partial L_y} = w/\left( \frac{dC_x/dQ_x} \right) \tag{3f.1}
\]

and

\[
\frac{\partial Q_y}{\partial L_y} = w/\left( \frac{dC_y/dQ_y} \right) \tag{3f.2}
\]

Dividing (3f.2) by (3f.1) gives:

\[
\left( \frac{\partial Q_y/\partial L_y} \right) \left/ \left( \frac{\partial Q_x/\partial L_x} \right) \right. = \left( \frac{dC_x/dQ_x} \right) \left/ \left( \frac{dC_y/dQ_y} \right) \right. \tag{3f.3}
\]

Both since

\[
\frac{dQ_y/dQ_x} = \left( \frac{\partial Q_y/\partial L_y} \right) \left/ \left( \frac{\partial Q_x/\partial L_x} \right) \right. \tag{3f.4}
\]

as shown in Appendix 2 below; it follows that

\[
\frac{dQ_y/dQ_x} = \left( \frac{dC_x/dQ_x} \right) \left/ \left( \frac{dC_y/dQ_y} \right) \right. \tag{3f.4'}
\]
FIGURE 3.1
the competitive equilibrium at point E, the effect of introducing monopoly here is to lower the output of \( Y \) and raise its relative price.

The situation will be aggravated if we now introduce elements of monopsony into the factor markets. Following on from the above quotation by Myint, for example, let \( Y \) operate as a dominant oligopsonist in both factor markets; that is, \( Y \) sets each factor price so as to equate its marginal revenue product with the corresponding marginal resource cost.\(^8\) This implies, of course, that \( Y \)'s marginal revenue product will exceed the corresponding equilibrium factor price; or

\[
(aQ_y/aL_y) \{ P_y(Q_y) + Q_y(dP_y/dQ_y) \} > w \tag{3.5}
\]

and

\[
(aQ_y/aK_y) \{ P_y(Q_y) + Q_y(dP_y/dQ_y) \} > r \tag{3.6}
\]

Similarly, if \( X \) is the subordinate competitor in both factor markets, then

\[
(aQ_x/aL_x) P_x = w \tag{3.7}
\]

and

\[
(aQ_x/aK_x) P_x = r \tag{3.8}
\]

Finally, on rearranging (3.5) through (3.8) it is clear that for most cases,\(^9\)

\[
(aQ_x/aL_x) / (aQ_x/aK_x) \neq (aQ_y/aL_y) / (aQ_y/aK_y) \tag{3.9}
\]

---

8. For a similar example, see Tisdell (1972, pp. 235-236). See also Leftwich (1973, ch. 18).

9. The exception is the case where the difference between sector \( Y \)'s marginal revenue product and the corresponding factor price is the same in both factor markets.
implying that an important efficiency condition for Pareto optimality is being violated. This situation is illustrated in Figure 3.1 by a point such as G, which lies inside the transformation curve, TT'. The effect of introducing monopsonistic conditions into the factor markets is thus to distort both the product and factor markets of our two-sector economy.

The policy measures relevant to the present analysis depend on whether the firms or industries concerned are 'natural' monopolies (and/or oligopolies). If they are not, then it should be theoretically possible to enforce a policy of marginal cost pricing in those markets characterized by imperfect competition; and should competitive conditions continue to exist elsewhere, then there is the chance that the policy would enable the economy to attain a Pareto-optimal allocation of resources. ¹⁰ Alternatively, if at least one firm or sector were a 'natural' monopolist or oligopolist, then the policy-maker would need to consider the theory of second best, according to which the (second best) solution to (Pareto) efficiency entails non-competitive pricing in all markets. ¹¹ Under our present assumptions, the relevant requirement is simply that the percentage divergence of price from the optimum should be the same everywhere: ¹² if monopoly existed in sector Y but perfect competition in X, for example, the optimal output would be that for which price exceeded


¹¹. See Samuelson (1947, pp. 252-253) and Lipsey and Lancaster (1956-57).

marginal cost by the same proportion in both sectors. Such a pricing policy should then enable the economy to achieve at least a constrained optimum: "The optimum solution finally attained may be termed a second best optimum because it is achieved subject to a constraint which by definition, prevents the attainment of a Paretian optimum".

Although the practical implementation of the latter policy is bound to run into numerous problems relating to identification and measurement, we can at least establish a general principle: namely, it is theoretically possible to achieve optimality by introducing price "corrections" designed to counteract those 'structural' distortions which cannot be eliminated directly; as Samuelson (1947, p. 252) might put it, under these circumstances "... two wrongs (do) make a right".

2. Factor Market Imperfections and Economic Dualism

It seems appropriate to begin an analysis of factor market imperfections with Arthur Lewis's (1954; 1972; 1979) model of economic dualism. The model generally views development as a relatively painless process according to which "surplus"
labour is transferred from a low-productivity or non-capitalist sector, to a high-productivity or capitalist sector. This transfer of labour is assumed to occur at a fixed wage, which enables capitalists to save, invest and continue employing surplus labour during the initial stage of development. Economic development is thus made synonymous with increased saving and investment: "... the central fact of economic development is capital accumulation".16

According to the dualist model, surplus labour in the non-capitalist sector is not openly unemployed in the conventional sense, but is rather 'disguisedly' unemployed or underemployed in the ("significant") sense that "... in many countries the market stalls are crowded with people who are not as fully occupied as they would wish to be".17 A useful interpretation of underdevelopment is that provided by Amartya Sen (1968) and A.P. Thirwall (1977), and illustrated in Figure 3.2 below. The northern portion of the vertical axis measures output in the non-capitalist sector, the east represents labour or the number of man- or labour-hours, and the south the number of labourers; TP, AP and MP are the total, average and marginal products of labour, respectively. Initially, let OS labourers each work OA/OS hours to produce the total output AQ. If OA/OR hours are,


17. Lewis (1972, p. 78). The question as to whether underemployment represents a situation of zero marginal productivity, has become one of the most contentious issues in development economics. And yet it is of very little consequence to the Lewis model: "Whether marginal product is zero or negligible is not of fundamental importance to our analysis" (Lewis, 1954, p. 419); indeed, in a subsequent paper Lewis (1972, p. 77) regretted every having raised the matter at all, claiming that "... this has merely led to an irrelevant and intemperate controversy".
Figure 3.2
however, assumed to represent the 'normal' working time per
labourer, then underemployment is simply defined equal to RS
labourers.\textsuperscript{18} This may of course imply that the marginal product
of \textit{labourers} is zero over some range beyond OA: for example, if
RS labourers were to leave the non-capitalist sector, then the
same output of AQ could be produced by OR labourers each working
OA/OR instead of OA/OS hours. But the marginal product of \textit{labour}
is positive and, in Figure 1, equal to the (hourly) wage rate.\textsuperscript{19}
Labour is employed up to the point where its marginal product
is equal to the wage rate, and underemployment takes the form
of a small number of hours worked per labourer. "It is not
that too much labour is being spent in the production process,
but that too many labourers are spending it."\textsuperscript{20}

On the above view, surplus labour is defined analytically
as simply \textit{unemployed} labour which, in the special circumstances
of developing countries, is 'disguised' or 'hidden' by a re-
distributive system of work (and income) sharing. The only
peculiarity of the dualist model is then the existence of a
fixed wage in the non-capitalist sector; but again, Lewis is
not concerned with the actual derivation of this wage: "The
model simply postulates as facts that in the initial stage the

\textsuperscript{18} See Thirlwall (1977, p. 104).

\textsuperscript{19} Alternatively, if the wage rate is assumed to equal the average
product in Figure 3,2, then the initial labour supply of OA hours
may be explained in terms of a work-leisure trade-off: beyond OA
labour, for example, the income per unit of time is not considered
worth the corresponding work effort; but when RS labourers leave
the non-capitalist sector, each of the remainder (i.e. OR labourers)
consider the additional income to be worth the extra effort. (See
also Lewis (1972); Thirlwall (1977); and Uppal (1969)).

\textsuperscript{20} Sen (1968, p. 3).
supply of labour at the given wage exceeds the demand, and that this condition will continue for some time despite the expansion of the capitalist sector.\textsuperscript{21} For expositional purposes, however, we shall continue to assume that the wage rate is equal to the marginal product of labour, thus implying the existence of a capitalistic wage-payment system in the non-capitalist sector.\textsuperscript{22}

The stage is now set for the emergence of Lewis's capitalist sector — which is presumably attributable to the advent of foreign investment during the "opening-up" process of economic development.\textsuperscript{23} In Figure 3.3 the non-capitalist and capitalist sectors are referred to as sectors X and Y respectively. Assume that for a given price of capital, wage rates are fixed at given minimum values in each sector, determined \textit{inter alia} by demographic, social and institutional factors.\textsuperscript{24} Assuming constant returns to scale in each sector, let the factor endowment in sector X be given by the input combination at J, representing $0_{xB}$ units of capital and $0_{xC}$ units of labour. Production in sector X initially occurs at E indicating employment of $0_{xB}$ capital and $0_{xA}$ labour units — assume at the minimum non-capitalist wage given by the slope of price line aa'. Surplus labour is thus equal to AC unemployed labour units and is, according to Lewis, available to sector Y at a (minimum) wage that exceeds the non-capitalist wage by some fixed proportion; e.g. 30 per cent, to cover the financial

\begin{itemize}
\item \textsuperscript{21} Lewis (1972, p. 77).
\item \textsuperscript{22} The same assumption is made by Fei and Ranis (1961) and Myint (1971).
\item \textsuperscript{23} See, for example, Myint (1971).
\item \textsuperscript{24} These factors are all mentioned in Lewis's original article (1954).
\end{itemize}
FIGURE 3.3
and psychological costs of a continuous labour transfer. If sector Y's wage is now given by the slope of price line \( w'w \) (passing through point E), then capital accumulation of \( O'yJ \) will be needed to induce (the equivalent of) AC labour units to migrate from sector X to sector Y. Production in Y will occur at E, indicating that \( O'yW \) labour and \( O'yJ \) capital units are being used to produce the output level, \( Y_5 \). Although the equilibrium combination at E represents full employment in the economy as a whole, it is clearly not Pareto-optimal in view of the presumed wage differential between sectors Y and X.

The amount of capital investment in sector Y depends on the prevailing wage differential between sectors Y and X and, more specifically, on the slope of price line \( w'w \) in Figure 3.3. It is generally acknowledged, however, that such wage differentials have widened considerably during most of the postwar period; indeed, labour in the capitalist sector tends to be relatively overpriced as a result of the particular wage policies pursued by post-colonial governments and trade unions. Moreover, capital is again relatively underpriced due to the existence of overvalued exchange rates, artificially low interest rates and various forms of tax exemption. In terms of Figure 3.3, the effect of such distortions \textit{ceteris paribus} will be to raise the slope of the price line in sector Y above that given by \( w'w \), thus

\begin{itemize}
\item[25.] Lewis (1954, p. 422).
\item[26.] See, for example, Taira (1966) and Tidrick (1975).
\item[27.] Todaro (1971).
\end{itemize}
further increasing the quantity of capital investment needed to attain overall full employment (i.e. in both sectors) at the input combination. But this would not change the general conclusion of Lewis's model in any essential way: that is, for any given wage differential between the capitalist and non-capitalist sectors, the cure for labour unemployment lies purely and simply with the ability of the former sector to accumulate capital.

The Lewis model is also not much altered if allowance is made for a change in the internal or external terms of trade. Lewis (1954) himself recognized, for example, that the growth of the capitalist sector may turn the terms of trade in favour of the non-capitalist sector. In terms of the above example, the implication of such a change in the terms of trade would be a reallocation of capital from the capitalist to the non-capitalist sector; and given the corresponding difference in the capital/labour ratio, it would also reduce the amount of capital investment needed to attain overall full employment. We shall return to this point below.

3. Economic Dualism and Rural-Urban Migration

It is possible to extend the basic Lewis model by considering some additional "imperfections" relating to the mobility of production factors. Our chief concern here is with Michael Todaro's (1969; 1971) celebrated model of rural-urban migration. The model represents perhaps the most significant challenge to

28. For a similar technical analysis, see Rybczynski (1955).
the then prevailing theories of labour migration: it is one of only a few studies that rejects the presumed two-dimensional relationship between labour migration and the corresponding wage differential(s) from the outset. Although there have been numerous tests of the ('classical') hypothesis of perfect factor mobility, it is probably fair to say that most of these studies constitute but slight variations on the basic Todaro theme.29

The primary aim of the model was to "explain the apparently paradoxical relationship of accelerated rural-urban migration in the context of rising urban unemployment".30 The model itself assumes that the migration decision depends not only on the relevant wage differential, but also on the probability of securing wage employment in the urban sector within a given time period. Specifically, if subscripts x and y refer to the rural and urban sectors respectively, then

\[ M_{xy} = f\left(\frac{w_y}{w_x}, \lambda_y\right) \]  

(3.10)

and

\[ \lambda_y = g(u_y) \]  

(3.11)

where \( M_{xy} \) is the number of labour units migrating from sector X to sector Y, \( w_y \) and \( w_x \) are the respective average wage rates, and \( \lambda_y \) is the (average) probability of obtaining wage employment in sector Y — assumed to vary inversely with the urban unemployment rate, \( u_y \). What this implies is that a given decrease in \( u_y \) is sufficient to encourage labour migration from sector X to


sector Y — notwithstanding the existence of a given wage differential, and irrespective of whether there are in fact new job openings available in sector Y. But any subsequent loss of income in sector X is more than likely to be offset in the minds of the migrants by the *expected* wage income in sector Y; as Todaro (1971, p. 411) puts it: "As long as the present value of the net stream of expected urban incomes over the migrant's planning horizon exceeds that of the expected rural income, the decision to migrate is justified".

The implication of Todaro's model for the dualist economy may be illustrated with the aid of Figure 3.4. Given *constant* returns to scale in each sector, let the output of X again occur at E indicating employment of $O_x A$ labour and $O_x B$ capital units; and similarly, let the remaining labour supply, namely $AC = O_y W$ units, be available for employment in sector Y. Suppose now that sector Y expands along the path $O_y E$, accumulating $O_y J$ capital units and employing $O_y W$ labour units, thus reducing $u_y$ to zero and encouraging additional labour to migrate from sector X to sector Y. In the absence of further capital accumulation, the effect of such migration will be to raise $u_y$ and lower $\lambda_y$ again, while at the same time causing a relative increase in the rural wage rate, $w_x$. Both these forces will continue until migration comes to an end at (say) point G, when $MA = WZ$ additional labour units have migrated from sector X to sector Y. These new migrants now become either unemployed or underemployed in sector Y where they are able to share in the work and income of those already employed; that is, the $WZ$ labour units continue to be available for employment.
FIGURE 3.4
at the ruling urban wage, $w_y$, and hence constitute surplus labour in the conventional sense of the word, as used in the Lewis model discussed above. This implies, of course, that the problem of labour underemployment has been simply transferred from the rural to the urban sector.

It is evident that further capital accumulation in sector $Y$ will simply raise $\lambda_y$ and induce yet more migration — notwithstanding the counteracting effect of subsequent increases in the rural wage, $w_x$. For illustrative purposes, one might imagine a situation where capital accumulation in sector $Y$ has eventually caught up with migration at a point such as $H$; that is, where the increase in $\lambda_y$ is just offset by the rising $w_x$ — shown here by the difference between the slopes of price line $nn'$ through point $H$, and the original price line, $aa'$ through $E$. But point $H$ also represents an increase in the capital stock equal to $O_yJ$ units in sector $Y$. Consequently, the implication of Todaro's analysis for the Lewis model is that it raises considerably the postulated amount of capital needed to provide job opportunities to the unemployed in sector $Y$.

Generally, the Todaro model suggests that policies aimed at providing employment in the urban areas require disproportionately large quantities of capital investment. This is due to the existence of factor price distortions generally, and the disequilibrating nature of rural-urban migration specifically, both of which tend to raise the amount of capital per unit of labour employed in the urban relative to rural sector. It follows that the cost, in terms of capital accumulation, of providing
additional employment opportunities is likely to be much smaller in the rural sector than it is in the urban sector. With reference to Figure 3.5 — which is a partial reproduction of Figure 3.4 — it is clear that capital accumulation in sector X would have to be accompanied by a corresponding wage increase sufficient to offset $\lambda_y$ and reduce $M_{xy}$ to zero; and as before, such a wage increase is given by the difference between the slope of price lines along the ray $O_XH$ extended, and the initial price line $aa'$ through point $E$. At a point such as $I$, for example, additional capital equal to $O'_yJ$ units are needed to attain full employment in the economy as a whole; but since some of this capital, namely $AA'$ units, has found its way to sector X, it follows that the total amount of additional capital required for full employment has been reduced from $O_yJ$ to $O'_yJ$ units.

The fact that this difference in capital accumulation is wholly attributable to the differences in factor prices and the capital/labour ratio between the urban and rural sectors, is obvious enough. Efforts to promote rural development should at least in principle raise the probability of obtaining rural rather than urban employment, thus limiting the incidence of rural-urban migration and further lowering the capital cost of attaining overall full employment.

4. The Urban Informal Sector

The urban informal sector has received a great deal of attention lately. Some observers have found it to be a major source of income and employment in the urban areas of developing
countries; 31 others have again been impressed by the competitive or unregulated state of its markets, its small-scale operations and use of labour-intensive production techniques, 32 while Marxists have denounced it as a "peripheral or marginal activity in the world system of capitalist production". 33 But despite these different emphases, few would deny that the urban informal sector owes its existence to the generally imperfect nature of the product and the factor markets in urban areas. 34 This relationship is partly reflected in the "illegal status" of informal sector activities, many of which do not as a rule comply with the licensing and other entry requirements imposed by such "formal" institutions as local government agencies, professional organizations, business associations and trade unions.

The origin of the urban informal sector may be related to the process of rural-urban migration according to Todaro. But instead of becoming unemployed or underemployed in the urban sector, newly arrived migrants are often able to raise capital and start new productive activities which are, on the whole, "... economically efficient, productive and creative". 35 Although the distinction between an underemployed person and informal sector employee may not be too apparent, it is nevertheless an analytically important one: whereas the underemployed are "... consumers of, but non-contributors to, the national product", 36 informal enterprises not only contribute to the

31. See, for example, Sethuraman (1977) and Souza and Tokman (1976).
32. ILO (1972).
33. Davies (1979, p. 91). See also Leys (1975).
34. See Truu and Black (1980).
35. ILO (1972, p. 51).
national product but they do so by using relatively labour-intensive production techniques. Such enterprises develop in response to prevailing market imperfections in the urban areas: for example, through surreptitious entry into an oligopolistic market, the informal entrepreneur is often able to supply additional units of the same or similar good or service at a comparatively low price; and similarly, by ignoring existing minimum wage legislation, the same entrepreneur may be able to employ available labour at a comparatively low wage. Whatever its legal status may be, the fact is that the urban informal sector "... attacks poverty directly by creating new sources of income and producing goods and services at low cost where they are needed most".

Informal sector activities play an important role in determining output and employment conditions in the urban sector generally. Referring to Figure 3.6, for example, assume that the rural sector, X, again produces at E using $O_A$ capital and $O_B$ labour units. Output in the urban sector, Y, first expands along the path $O_Y$, resulting in employment of $O_J$ capital and $O_W$ labour units, assume at a "formal" or institutional wage given by the slope of price line $ww'$. The accompanying decrease in $u_Y$ (or increase in $\lambda_Y$) encourages (additional) labour to migrate from sector X to sector Y, all of whom are now assumed to become employed in a newly emerging urban informal (sub-) sector, assume at a competitive wage which is below the institutional one. According to the Todaro model, the emergence

37. Although this is not usually reflected in official statistics.

FIGURE 3.6
of such an urban sub-sector may itself affect the migration decision of prospective migrants in the rural areas. There are two possibilities here. Firstly, if prospective migrants regard the informal sector as a potentially permanent work destination, then the growth of informal job opportunities may well lower $u_x$ and raise $\lambda_y$ again, inducing yet more labour to migrate from sector X to sector Y; but the resultant increase in urban labour supply should also reduce the urban informal wage, and hence the average urban wage ($w_y$) too, thus eventually bringing migration to an end when, say, $MA = WZ$ labour units have left sector X to join the urban informal sector.

Secondly, if the informal sector is regarded simply as a temporary transit to better paid jobs in the rest of sector Y, then its effect on migration is essentially the same as that of a pool of unemployed or underemployed urban dwellers; that is, the prospective migrant views the urban informal sector merely as a manifestation of unemployment --- albeit a relatively gainful one --- and continues to regard the institutional urban wage and related job probability in the modern or formal sector as the only relevant determinants of migration. Accordingly, migration will proceed until $u_y$ has increased sufficiently to reduce $\lambda_y$ to its original level; assume, for example, that this also occurs at G, that is, when $MA = WZ$ labour units have migrated from sector X to sector Y.

In each of the above cases, the output and employment effects work in the same general direction. In Figure 3.6, suppose that production functions are identical in informal enterprises and the
rest of sector Y, but that the former use more labour relative to capital due to the corresponding wage differential. The net result is then that the decrease in the average urban wage, $w_y$, also lowers the average capital/labour ratio in sector Y, enabling it to produce at a point such as G, representing production of the composite output level $Y_6$ by means of $O''J_y$ capital and $O_Z$ labour units. This means that the growth of informal sector activities effectively lowers the amount of capital investment needed to achieve full employment in the economy as a whole, e.g. $O''J_y$ compared to $O'yJ$ capital units in Figure 3.6; indeed, the effect is similar to that of a labour subsidy in the sense that they "get the prices right," or at least lessen the relative importance of factor price distortions in the urban sector as a whole.

5. General Policy Implications

The analysis thus far has focused on the attainment of full employment through a net increase in the capital stock. For policy purposes, however, it is important also to consider the means by which existing supplies of capital and labour can be used to overcome the adverse employment effects of market imperfections. It seems therefore worthwhile to assess the general desirability of using indirect taxes and subsidies for this purpose; or more precisely, of using the particular tax-subsidy combinations suggested by both the Todaro hypothesis and the above analysis of the urban informal sector.
To begin with, let us consider the following general case of the market imperfections hypothesis. Suppose that variations in the prices of capital and labour are institutionally constrained by given maximum and minimum values respectively, in both the urban and rural sectors of a typical developing economy. It seems reasonable also to assume that the maximum price of capital in the rural sector, X, exceeds that in sector Y, while the minimum wage in sector Y again exceeds that in sector X. This is illustrated with the aid of Figure 3.7(a), where the solid line O'R' depicts the locus of price lines for which the price of capital in sector X is at a maximum and the wage at a minimum; and similarly O'S' gives the locus of sector Y's price lines for which the corresponding price of capital is at a maximum and the wage at a minimum. The difference in the slope between O'R' and O'S' is due to the abovementioned differences in these maxima and minima. Since this implies that the slope of the price line in each sector is only permitted to increase with respect to price lines along O'R' and O'S' respectively, it follows that the area O'C'O'R' represents the feasible (economic) region of production for sector X, while O'CS' is the corresponding region for sector Y. Similarly, the "efficiency locus" or path of maximum total output, is given by the line segments O'H' cum S'H' and H'O', which may now be used to trace out the corresponding (constrained) transformation curve. This is shown as the curve SHO_y in Figure 3.7(b), where the segments SH and HO_y correspond to the similarly lettered portions of the efficiency locus in Figure 3.7(a).

39. Although the price of capital tends to be relatively under-priced in both the rural and urban sectors (Todaro (1971)), it is nevertheless higher in the rural sector (Chenery et al. (1974)).
Output combinations along $HO_y$ represent full employment in the economy as a whole, while the output combinations given by $SH$ indicate labour unemployment.

Referring back to Figure 3.7(a), suppose that production in sectors X and Y occurs at $E'$ and $F'$, respectively, indicating labour unemployment equal to $A'D'$ or $X'Z'$ units. These input ratios correspond to the equilibrium output ratio $\alpha$ in Figure 3.7(b), where the community indifference curve labelled $W_1$ is tangent to the commodity price line $P_1P_1'$. The slope of $P_1P_1'$ is evidently less than that of the constrained transformation curve, $SHO_y$, which indicates simply that the wage (and interest) differential is assumed to be "working against" sector Y. 40

The solution to this unemployment problem may, of course, include virtually any output combination along the contract curve, $Q'O_{xy}'$ in Figure 3.7(a). But it nevertheless seems reasonable to assume that a policy aimed at expanding output of both commodities would be deemed preferable to one involving a decrease in either of the two commodities. Such a policy should allow production of X and Y to move to a desired output combination on the contract curve, anywhere between the points where the latter curve intersects the respective isoquants, $X_3$ and $Y_1$; e.g. the combination at point $B'$, where the respective marginal rates of technical substitution of labour for capital are equal. The net effect would then be a movement from $\alpha$ on

---

40. See, for example, Hagen (1958), Bhagwati and Ramaswami (1963), Johnson (1965), Battra and Pattanaik (1970), Batra (1973, ch. 10) and Hazari (1978, Parts 1 and 2).
the constrained transformation curve, to \( \beta \) on the true transformation curve in Figure 3.7(b), which *ceteris paribus* implies achievement of the higher level of social welfare associated with indifference curve \( W_2 \).

The desired increase in sector Y's output clearly requires the use of a labour subsidy and/or capital tax — the effect of which would be broadly similar to that of promoting informal sector activities at the expense of production in the rest of sector Y. In the absence of a cost constraint, the labour subsidy (or capital tax) would not only lower the price of labour relative to that of capital, but would also, given variable technical coefficients of production, enable sector Y to expand production through the use of relatively labour-intensive techniques. Since the actual wage accruing to the employee is unlikely to be affected by the subsidy itself, employment growth in sector Y may well raise the probability of obtaining employment (\( \lambda_y \)), and hence encourage (potentially unemployed) labour to migrate from sector X to sector Y. According to Todaro's model, the only way to prevent such migration would be to raise the actual wage in sector X sufficiently so as to counteract the rising \( \lambda_y \).41

Moreover, the increase in the rural wage rate would almost certainly have to be supplemented by an appropriate factor tax-subsidy mix: for example, if the *required* wage increase were such as to raise the slope of sector X's price line above that

41. Assuming, of course, that a decrease in the urban wage rate is not permitted.
given by $tt'$ — in which case production of $X$ would occur somewhere to the left of $e'$ — then a labour subsidy and/or capital tax would be needed to bring production of $X$ back to the optimal input combination at $e'$; or alternatively, if the effect of the wage increase were to raise the slope of sector $X$'s price line to below that of $tt'$, then the appropriate policy measure would be a labour tax and/or capital subsidy. It thus appears that the size and composition of the optimal factor tax-subsidy package in each sector depends on (i) the desired increase in the rural relative to urban sector's output, and (ii) the extent to which the rural wage rate would have to rise in order to remove the (Todarian) migration incentive.

Consequently, if Figure 3.7 does indeed represent an approximation of conditions in developing countries, then it follows that there are at least two conditions for eliminating unemployment in the face of a capital constraint. The first and necessary condition is an increase in output in both the rural and urban sectors, by means of factor rather than product subsidies or taxes — which seems consistent with the proven superiority of a factor subsidy over both a product subsidy and tariff in the elimination of international trade disequilibria arising from factor price distortions. But the sufficient condition is rather less conventional, in that it entails an increase in the rural wage rate sufficient to offset the relatively high probability of obtaining employment in the urban sector.

42. See Chapter 4 below.
1. Introduction: The Gains from Trade

The basic reason why countries are said to engage in international trade is that it enables them to consume more than they can actually produce with their given resource endowments. This statement is, of course, a direct corollary to the Ricardian theory of comparative advantage. According to Ricardo, a country stands to gain if it specializes in producing those commodities in which it has a so-called comparative cost advantage, and exchange them internationally for commodities in which it has a comparative disadvantage. If all countries were to produce and trade in this manner, the value of world production and consumption would be higher than it would be in the absence of international trade.

In a two-country, two-good world, for example, suppose country A has a comparative advantage with respect to good X insofar as she can produce it relatively more cheaply than country B; that is, country A foregoes fewer units of good Y in order to produce one unit of X, than does country B. Now, if country B has a comparative advantage in the production of good Y, then it follows that country A can get more of Y in exchange for X internationally than she can do at home. Compared to the pre-trade situation, this means that free trade will bring

1. A good summary of the theory is provided by Södersten (1970, ch.1).
about an increase in the relative price of good X to country A, while at the same time raising the relative price of good Y for country B. Accordingly, if country A exports good X and imports good Y from country B, then both countries benefit in that each will be able to obtain the other's good more cheaply than if these were produced domestically: "... the gain from trade is the difference between the value of things that are got and the value of things that are given up".2

The static gains from trade may be illustrated with the aid of Figure 4.1. The non-trade or autarkic equilibrium in country A occurs at point E, where the community indifference curve labelled \( W_1 \) is tangent both to the transformation curve \( TT' \), and to the domestic price line \( d_1d_1' \). The international or external terms of trade are in turn given by price line \( f_1f_1' \), indicating a higher relative price for good X on the international market. The (domestic) production point accordingly moves from E to F along \( TT' \), where the marginal rate of product transformation is equal to the corresponding international price ratio — given by the slope of \( f_1f_1' \). The country may now export AF units of good X in exchange for AG units of good Y, and thus consume at point G where the community indifference curve labelled \( W_2 \) is tangent to price line \( f_1f_1' \). While Pareto optimality is assured by the fact that the community marginal rate of substitution equals the marginal rate of product transformation, it is clear that

specialization and free trade have enabled the country to achieve a higher level of social welfare than enjoyed previously; as indicated by $W_2$ versus $W_1$.

Turning now to the dynamic gains from trade, it is perhaps worth noting that these too had been first recognised and considered at some length by the classical economists. More recently, Myint (1954-55; 1958; 1968), Haberler (1959) and others have argued that the "indirect" or dynamic benefits from trade may be at least as important as the static gains highlighted in the traditional theory of comparative advantage: "If we were to estimate the contribution of international trade to economic development, especially of the underdeveloped economies, solely by the static gains from trade in any given year on the usual assumption of given production capabilities, we would indeed grossly underrate the importance of trade".

The chief dynamic gain from trade is the increase in the size of the total market available to a trading country's producers. The discovery and subsequent utilization of new export markets generally give rise to economies of scale, an inflow of capital and other productive resources and the international dissemination of technical knowledge. Such benefits amount to an increase in the physical quantity and/or productivity of resources used in the production of both exports and import-competing commodities; or as Haberler (1959, p.108)

---

4. See Hicks (1953); Cairncross (1960; 1962) and Nourse (1961).
puts it, the dynamic gains from trade represent "an outward shift of the production possibility curve brought about by a trade-induced movement along the curve". However, if the value of the country's trade is large in relation to that of its trading partners, then the increased supply of its exports, coupled with the increased demand for imported commodities, could raise the relative prices of its imports on the international market\(^6\); and to the extent that this happens, the resulting deterioration of its terms of trade should at least partly offset the dynamic gains from trade.

All this may be illustrated with the aid of Figure 4.2. Domestic production and consumption initially occur at points F and G, respectively, indicating equality between the community marginal rate of substitution, the marginal rate of product transformation and the international price ratio as given by the slope of price line \(f_1f_1'\). Suppose now that trade induces export-biased growth\(^7\) resulting specifically from (i) an increase in the supplies of capital and labour, and (ii) (neutral) technological progress in the export sector, \(X\). Such a situation is shown by the non-parallel, outward shift of the transformation curve from \(TT'\) to \(SS'\)\(^8\). If the international terms of trade remain unchanged, the production point will move from F to H, and the consumption point from G to I, with the social indifference curve labelled \(W_3\) tangent to the international price line \(f_2f_2'\).

---

6. The effect of international trade on the terms of trade was first recognised by several English economists during the previous century. (See Södersten, (1970, pp. 184-190)).
7. See, for example, Caves and Jones (1973, chs 25 and 26).
8. On the nature of this shift, see Krauss and Johnson (1974 and Baldry (1980, ch. 6)).
(parallel to $f_1f'_1$). Under the large-country assumption made above, it is possible that the (trade-induced) growth could cause an increase in the relative price of the imported commodity $Y$ — illustrated here by the indifference between the slope of the new price line, $g_1g'_1$, and the slope of $f_2f'_2$. Although the latter change in the terms of trade, ceteris paribus, lowers the level of social welfare (for example, from $W_3$ to $W_2$), the net effect is nevertheless an increase in the welfare level from $W_1$ to $W_2$.

The question naturally arises whether the welfare effect of the thus deteriorating terms of trade could in fact outweigh the dynamic (welfare) gain from trade; that is, whether the deteriorating terms of trade could bring about a net decrease in the social welfare level. This possibility has been frequently raised in the context of developing countries, for example by such writers as Singer (1950), Prebisch (1950; 1959), Viner (1953), Myint (1954-55) and Myrdal (1956), and in Bhagwati's (1958) rather elegant interpretation of "immiserizing growth". In what follows, therefore, we shall consider the various reasons and policy proposals put forward for the alleged deterioration of the developing countries' terms of trade.

2. The Prebisch-Singer Thesis.⁹

Prebisch divides the world into industrial or developed 'centres' which generally specialize in the production of manufactures; and 'peripheral' or developing countries

⁹ See Baer (1962) and Flanders (1964).
specializing in the production of agricultural and other primary commodities. His main contention is that during most of the past 100 years the terms of trade have turned in favour of centre countries and against the periphery: 
".... the price relation turned steadily against primary production from the 1870's until the Second World War". 10

Prebisch bases his explanation for the periphery's deteriorating terms of trade on two critically important assumptions, namely, (i) while the income elasticity of the periphery's demand for the exports of the centre exceeds unity, the corresponding elasticity of the centre's demand for the periphery's exports is again significantly smaller than one; and (ii) the commodity and factor markets in the centre are generally characterized by imperfect competition, while those in the periphery tend to approximate perfect competition.

The process begins with a given technological innovation giving rise to productivity increases in the export sectors of both the centre and the periphery countries. Under competitive market conditions, such productivity increases are likely to be accompanied by corresponding price reductions which occur as a result of new firms entering the relevant product markets; similarly, if the rate of productivity growth is the same everywhere, the "benefits of technological progress (will) tend to be distributed alike over the whole community". 11 But since productivity gains in the centre have

in fact been much larger than those in the periphery, the latter reasoning would lead one to expect that the terms of trade should have turned in favour of the peripheral countries. This has evidently not happened. As Prebisch and Singer see it, the main reason for the periphery's declining terms of trade should be sought in the imperfectly competitive nature of the centre's product and factor markets. Any given productivity increase in the centre will either encourage trade unions to bid up wages in an attempt to raise the real income of labour; or alternatively, if trade unions are unable to enforce such wage increases, prices are unlikely to fall in view of the imperfectly competitive structure of the product markets. Likewise, the relatively high income elasticity of the periphery's demand for the centre's exports is bound to put a further upward pressure on the latter's export prices, or at least prevent them from falling.

As far as the peripheral countries are concerned, productivity increases tend to be matched by corresponding price reductions. Since technological innovations in the periphery often entail a large-scale substitution of capital for labour, Prebisch maintains that the level of labour employment may actually fall in the absence of a sufficiently large increase in aggregate output, thus lowering wages or at least leaving them unchanged at subsistence level. Any profits thus made will, however, soon disappear as prices are bid down by new

12. See, for example, Singer (1950).

firms entering the product markets of the periphery. Similarly, this downward trend in prices will be reinforced generally by the fact that the income elasticity of the centre's demand for peripheral exports is relatively small. It therefore follows that while the periphery's export prices tend to fall in the face of given increases in productivity, those of the centre remain unchanged and may even rise: as Prebisch puts it, "... the great industrial centres not only keep for themselves the benefit of the use of new techniques in their own economy, but are in a favourable position to obtain a share of that deriving from the technical progress of the periphery". 14

A diagrammatic illustration of the Prebisch-Singer thesis is provided by Bhagwati's (1958) notion of immiserizing growth. In Figure 4.3, for example, we reproduce the growth-induced, outward shift of the transformation curve shown in Figure 4.2. For reasons discussed earlier, suppose the latter shift is accompanied by a relatively large increase in the relative price of the imported commodity, Y, for example to the level given by the slope of the new international price line $h_1'h_1'$. Production accordingly adjusts to point U while, at the new terms of trade, consumption occurs at point V. But since U represents a lower level of social welfare than the original, pre-growth equilibrium at G, it follows that the country has been made worse off by the growth-induced decline in its terms of trade: "Economic expansion increases output which, however,

FIGURE 4.3
might lead to a sufficient deterioration in the terms of trade to offset the beneficial effect of expansion and reduce the real income of the growing country.\textsuperscript{15}

It is, finally, possible to extend the above analysis by considering two recent studies undertaken by Bhagwati and Brecher (1980) and Brecher and Choudhri (1980). These authors contend that the dynamic gains from trade would be limited — or the welfare loss increased — if capital investment were undertaken by foreigners who received the whole or a part of the value of the marginal product of capital. For this purpose, it would be necessary to distinguish between an 'aggregate' and a 'national' transformation curve (and budget line), where the former refers to the total value of production, and the latter to the total value net of the returns accruing to foreign-owned capital. In terms of Figure 4.3, for example, the appropriate 'national' transformation curve and price line would lie somewhere to the left of $SS'$ and $h_1h_1'$ respectively, thus further lowering the welfare level obtainable under conditions of immiserizing growth. The latter conclusion seems at least consistent with Singer's (1950) view that foreign investment convey very few benefits to developing countries on the whole, and may even make them worse off via a deterioration in their terms of trade.

Prebisch and Singer are perhaps best known generally for their respective policy proposals. Prebisch in particular advocates a policy of tariff protection coupled with selective

\textsuperscript{15} Bhagwati (1958, p. 325).
export promotion, as a means of arresting the periphery's declining terms of trade, and possibly even of improving them over time. Singer in turn invokes the classical infant-industry argument by emphasizing the dynamic benefits that can be gained from a policy of tariff protection. Since these policy measures have evoked much lively discussion in the literature lately, it seems worth analyzing them here explicitly.

3. Tariff Protection and the Terms of Trade

The terms of trade (or "optimum tariff") argument for protection depends on whether the tariff is likely to turn the international terms of trade in favour of the tariff-levying country. If the country is, for example, too "small" to affect world prices, then the domestic price of its imports will rise by the full amount of the tariff, in which case the country is bound to experience a deteriorating welfare position. This may be illustrated with the aid of Figure 4.4 below, where the initial production and consumption equilibria are shown at points F and G, respectively. Suppose a tariff is now levied on the imported commodity, Y, so that its relative price increases from that given by the slope of the international price line $f_1f_1'$, to that given by the slope of the (domestic) price line, $d_1d_1'$. Given maximizing behaviour on the part of producers and consumers alike, the former will respond by


17. See Metzler (1949), Södersten and Vind (1968) and Johnson (1964). Good summaries of the literature can be found in Corden (1971) and Caves and Jones (1973, ch. 12).
FIGURE 4.4
producing more of good Y relative to X, while the latter will again consume more of good X relative to Y, thus establishing new equilibria at points H and I, respectively; that is, the economy will now be exporting AH units of good X in exchange for AI units of good Y. At the tariff-inclusive price of good Y, consumers will exchange AH units of good X for AB units of Y, leaving the remaining quantity of BI units to be collected by the government in the form of tax revenue. If the latter amount is returned to consumers in the form of a lump-sum payment for spending purposes, however, consumption will occur at point I, where the community indifference curve labelled W₁ is both tangent to the domestic price line d₂d₂' (parallel to d₁d₁'), and intersects the international price line, f₂f₂'.

This example illustrates how the imposition of a tariff could distort the international economy: although both the community marginal rate of substitution and the marginal rate of product transformation are equal to the domestic price ratio, the latter is nevertheless less than the international price ratio — given here by the slope of price line f₂f₂'. This distortional effect is further reflected in the fact that the new equilibrium at point I represents a lower welfare level than the initial, free-trade equilibrium at G. It is important to note, however, that the latter welfare loss is due partly to the small-country assumption made at the outset: "... the optimal tariff is zero if the terms of trade cannot be altered".¹⁸

The situation will be different if the value of the
tariff-levying country's trade is large relative to that
of the rest of the world. Here it would at least be
possible for the country to improve its terms of trade,
and to achieve a welfare gain, via an appropriate policy
of tariff protection. In Figure 4.5 below, for example,
we reproduce the initial free-trade equilibria shown in
Figure 4.4: domestic production and consumption occur
at the respective points F and G along the international
price line, \( f_1f_1' \), indicating attainment of the social
welfare level, \( W_2 \). The immediate effect of the imposition
of a tariff, *ceteris paribus*, is to raise the relative
price of the imported commodity \( Y \). Suppose, however, that
the accompanying decrease in the country's import demand
is sufficient to lower the *international* price of good \( Y \)
to the level given by the *newly* established international
price line \( g_1g_1' \). The latter price decrease implies, of
course, a correspondingly lower tariff-inclusive *domestic*
price of good \( Y \) — shown here by the slope of the
parallel price lines, \( e_1e_1' \) and \( e_2e_2' \). The consumption
equilibrium point accordingly moves from G to M on the
new international price line \( g_1g_1' \), where the community
indifference curve labelled \( W_3 \) is tangent to the domestic
price line, \( e_2e_2' \). But since point M represents a higher
welfare level than the initial position at G, it follows
that the tariff has indeed made the country better off
than before; in short, the tariff-induced improvement in
the terms of trade has more than offset the distortional
effect of the tariff itself.
FIGURE 4.5
It is probably fair to say that Prebisch based his argument for tariff protection on the large-country assumption made in the previous paragraph: throughout his two major works, for example, he implicitly assumed the existence of a two-country world consisting of one centre and one peripheral country. But even if the peripheral country did act as a monopolist with respect to the centre, it seems doubtful indeed whether a uniform policy of tariff protection would enable the periphery to improve its terms of trade to any significant extent. The fact of the matter is that the value of peripheral imports from the centre constitutes but a negligibly small proportion of the value of total output of the centre countries: "... the centre's biggest customer is itself, not the periphery". Moreover, export prices in the centre are unlikely to fall in response to a significant decrease in demand, in view of the abovementioned downward inflexibility of prices and wages in the centre's product and factor markets; June Flanders (1964, p. 302) puts it succinctly: "... a downward shift in the periphery's demand function for centre exports will result in making the centre worse off, through unemployment, without making the periphery better off through an improvement in the terms of trade".

4. The Infant-Industry Argument for Protection

Tariff policies in developing countries generally attempt to protect domestic infant industries against foreign competition, with a view to creating more job opportunities for a growing population. The infant-industry case for protection is thus largely a dynamic one and, as such, similar to the dynamic argument for free trade; specifically, a tariff-induced increase in the (domestic) production of importables usually gives rise to an inflow of capital, technical know-how and other resources, while at the same time enabling domestic producers to reap the benefits of various internal and external economies of scale. This situation may be illustrated with the aid of Figure 4.6 below, where the initial production and consumption equilibria are shown at points F and G, respectively. The imposition of a tariff on good Y, ceteris paribus, raises its relative price to that given by the slope of the tariff-inclusive domestic price line, \( d_1d_1' \); and as before, the production and consumption points move to H and I respectively, indicating a decrease in the level of social welfare from \( W_2 \) to \( W_0 \).

Now, it is possible to show that the latter welfare loss may be more than offset by growth-induced increases in the level of real income. Suppose, for example, the dynamic gains from protection are biased in favour of the import-competing good, Y, so that the transformation curve undergoes

---

20. See, for example, Singer (1950), Myint (1963), Södersten (1970, ch. 21) and Corden (1974). In addition, Colman and Nixson (1978, ch. 8) provide a useful summary of the nature and general economic consequences of "import-substituting" strategies in various developing countries.
a non-parallel, outward shift from TT' to SS'. If the country now removed the tariff altogether without affecting the international terms of trade, it would produce at point J and consume at M, thus attaining the higher welfare level given by W₃.

Whether the dynamic gains from tariff protection will exceed the corresponding gains from free trade is, of course, an open question. But even if it did, this would not in itself provide a justification for tariff protection. It is generally agreed that while the same objective(s) could be achieved through the use of domestic subsidies and taxes, such policies would nevertheless entail a smaller sacrifice in terms of real income foregone, than would a policy of tariff protection. Suppose, for example, the government believed that a given increase in the domestic production of importables is justified on account of the dynamic benefits associated with such an expansion. Specifically, let the desired level of production be AH units of good Y in Figure 4.6. We have already shown that the imposition of a tariff on good Y could shift resources from the initial, free-trade equilibrium, F, to the desired allocation at point H. Alternatively, the government could subsidize the domestic

21. This would depend on the size and nature of the respective dynamic benefits, as well as on the size of the country’s trade relative to that of the rest of the world.

22. See Corden (1957), Johnson (1965) and Bhagwati (1971).

23. Notwithstanding, that is, the short term loss of welfare.
production of good Y, and/or tax that of good X, in order to effect the same movement of the (equilibrium) production point along the transformation curve, TT'. The difference is simply that consumers would now be allowed to purchase commodities at existing world prices, indicated here by the slope of the international price line, $f_3'f_3$; similarly, the new consumption equilibrium would occur at point N, where the community indifference curve labelled $W_1$ is tangent to price line $f_3'f_3$. Consequently, while the dynamic benefits would presumably be the same for both policies, it is clear that the welfare loss resulting from the tax-cum-subsidy on domestic production is smaller than the corresponding loss associated with tariff protection.

5. **Market Imperfections and Trade Policy**\(^{24}\)

An import tariff is not the only type of distortion affecting the international trade relations between countries. There are many others too that may drive a wedge between a commodity price and its marginal social cost of production, and/or the corresponding marginal social benefit. Perhaps the most common of these are external economies arising from the interdependencies of consumption, production and investment decisions, the existence of monopolistic and oligopolistic markets and various policies of wage determination. In all such cases domestic prices do not as a rule reflect their true marginal cost of transformation in production.

---

From a development point of view, it is interesting to compare the various policy options open to the government when faced with a monopolist controlling the industry in which the country has a comparative advantage. There are essentially two possibilities here. The first is illustrated with the aid of Figure 4.7 below, which shows the situation for a small country specializing in the production of good X. If X is controlled by a monopolist while good Y is produced by one of a large number of small-sized competitors, the autarkic equilibrium will occur at a point such as E, where the community indifference curve labelled W₁ is tangent to the domestic price line, d₁d₁'; the fact that the community marginal rate of substitution exceeds the marginal rate of product transformation, indicates simply that the domestic price of good X is higher than its marginal cost of production. Notwithstanding the latter discrepancy, we assume here that good X is relatively more expensive on the international market than it is at home — illustrated, for example, by the difference between the slope of the international price line, f₁f₁', and the slope of price line d₁d₁'. This implies that under free trade domestic production will adjust to a point such as F on the transformation curve TT', and consumption to point G, thus enabling the economy to attain the higher level of social welfare W₂.

Although free trade is evidently deemed preferable to autarky in the example represented by Figure 4.7, the former nevertheless represents a sub-optimal allocation of resources due to the continued existence of monopolistic conditions in
FIGURE 4.7
industry X. Bhagwati and Ramaswami (1963) have shown that the optimal policy here would be a tax-cum-subsidy aimed directly at the source of the distortion, namely domestic production: specifically a subsidy on good X (and/or tax on good Y) could move the equilibrium production point from F to H along the transformation curve TT', and the corresponding consumption point from G to I, and thus eliminate the divergence between the marginal rate of product transformation and the corresponding international price ratio. Such a policy would thus enable the economy to achieve the highest welfare level shown in Figure 4.7, namely W₃.

The second possibility referred to above was raised by Gottfried Haberler in a celebrated article which appeared in the Economic Journal in 1950. He proved that autarky could yield a better solution than free trade if domestic prices were significantly different from their respective marginal (social) costs of production. This situation is shown in Figure 4.8. The initial equilibrium under autarky occurs at point E on the transformation curve, TT', indicating a relatively substantial divergence between the community marginal rate of substitution and the marginal rate of product transformation for good X. If the magnitude of the distortion is such as to raise the relative domestic price of good X above its corresponding international level, then free trade will cause the production point to move from E to M, and the consumption point from E to N, where the indifference curve labelled W₀ is tangent to the international price line,
FIGURE 4.8
The free trade solution is thus deemed inferior to the autarkic equilibrium.

The above example illustrates how free trade could aggravate the effect of the domestic distortion by encouraging a shift of resources away from the industry in which the comparative advantage lies; in point of fact, at the free-trade equilibria given by points M and N, the country would be exporting good Y instead of good X. Although an appropriate policy of tariff protection could give rise to a welfare improvement, the optimal policy again entails a tax-cum-subsidy on domestic production. Such a policy would eliminate the effect of the distortion entirely, by shifting resources from point E to point P along the transformation curve, thus enabling the economy to maximize welfare at the output mix, Q.

Turning, finally, to the policy implications of factor market imperfections in an open economy, consider the situation set out in Figure 4.9. The constrained transformation curve, SAT', is similar to that shown in Figure 3.7(b): it embodies the fact that while production is constrained by the existence of minimum wage rates throughout the economy, the minimum wage is nevertheless higher in sector Y than in sector X. The latter discrepancy also accounts for the fact that at the initial autarkic equilibrium, E, the marginal rate of product transformation exceeds the slope of the domestic price line, $d_1d_1'$.\(^{26}\)

---

25. See Bhagwati and Ramaswami (1963).

26. It will be recalled that point E represents a situation of labour unemployment in the economy as a whole. (See chapter 3, section 5).
Suppose now that free trade raises the relative price of good X to that given by the slope of the international price line, $f_1 f_1'$. Domestic production accordingly adjusts to point Q, and consumption to point R, thus indicating attainment of the higher level of social welfare, $W_2$. 27

The fact that the free trade solution is not necessarily superior to autarky in the presence of factor price distortions, need not concern us here: 28 the general policy implications are the same irrespective of whether free trade is preferred to autarky. A tax-cum-subsidy on domestic production, for example, would be preferable to free trade if it caused a shift of resources to, say, point A on the constrained transformation curve, in which case the new consumption point would presumably lie somewhere to the northeast of point R. But such a policy would be sub-optimal insofar as it would not be directed at the source of the domestic distortion, namely the labour market.

What is needed here is a policy of tax-cum-subsidy on factor use that would enable the economy to produce along its true transformation curve, $TT'$ — for example, at point V where the marginal rate of product transformation equals the slope of the international price line, $f_2 f_2'$. The economy would then maximize welfare by consuming at point Z, indicating equality between the community marginal rate of substitution, the marginal rate of product transformation and the corresponding international price ratio.

27. For similar analyses, see the references listed in footnote 24 above.

28. It can be shown that if (i) the wage differential "worked against" sector X instead of sector Y, and (ii) the slope value of the domestic price line exceeded that of the international price...
More generally, there can be little doubt that the above analysis is particularly relevant to developing countries whose product and/or factor markets are characterized by various price distortions. Our chief conclusion is simply that for a small country at least, free trade is generally superior to a policy of tariff protection. In the presence of domestic price distortions, however, both tariff protection and free trade compare unfavourably with domestic subsidies and taxes aimed specifically at the source of the distortion, whether it be the product or the labour market. It is indeed worth quoting Meier's (1976, p. 652) summary of the policy measures relevant to factor price distortions:

"The difficulty with protection by a tariff is that it seeks to remedy the distortion by affecting foreign trade, whereas the distortion is in a domestic factor market. In this case, a policy of subsidization of production of the import-competing commodity, or of taxation of agricultural production, would be superior to a tariff. A policy of subsidization on the use of labour in the import-competing industry, or a tax on its use in agriculture, would be an even better solution; since it directly eliminates the wage differential, this policy yields a higher real income than can be attained by a tax-cum-subsidy on domestic production. A tariff on industrial imports is thus the least effective way of offsetting a distortion in a labour market."

line, then the free trade equilibrium would be deemed inferior to that associated with autarky. (See also Battra and Pattanaik (1970)).
CHAPTER 5

INCOME REDISTRIBUTION, EMPLOYMENT AND ECONOMIC GROWTH

1. Introduction

The previous chapters have deliberately ignored the general effect of economic development on the personal distribution of income. This was done largely for expositional purposes. It could perhaps be argued that the various explanations put forward for the emergence and continued existence of labour unemployment, did at least implicitly recognize the potentially inequitable consequences of particular development processes.1 Nevertheless, in the two-sector model proper — that is, in the absence of a government or any other redistributive institution — an unemployed person could not earn any income and would, presumably, succumb to the traditional Malthusian constraint.

The relationship between economic development and income distribution has attracted much attention in the recent literature. It came to light in the well-known study of the Kenyan economy by the International Labor Office (IL0, 1972), which was subsequently followed by several similar works, both of a theoretical and an empirical nature;2 indeed, from the point of view of development economics at least, the seventies may be considered the decade of "redistribution with growth".3

---

1. The terms "inequitable" and "equitable" are used here in the general Pigovian sense: any given decrease in the share of real income accruing to the poor will, ceteris paribus, lead to a decrease in the level of social welfare. (See Pigou (1960, ch. 8)).
2. Two recent surveys are Morawetz (1974) and Cline (1975).
3. After Chenery et. al. (1974).
Contrary to the conventional (i.e. Kaldorian) wisdom, the ILO maintained that most developing countries were either able to achieve the twin objectives of economic growth and an equitable (or more equal) distribution of income simultaneously, or that they could at least achieve the latter at a relatively small cost in terms of output foregone. This view has not gone unchallenged, however, as several writers have found that for some countries at least, a redistribution of income in favour of the poor might well involve a significant trade-off in the growth of both output and employment. 4

In section 2 below we consider the ILO argument in the context of the two-sector models developed in the earlier chapters. Section 3 invokes the so-called demonstration effect in an attempt to show how changes in individual tastes could limit the employment potential of a policy of income redistribution. Finally, section 4 discusses the implications of income redistribution for economic growth.

2. Income Redistribution and Employment: The ILO View

Traditionally the distribution of income is held to be an important determinant of the savings ratio, and hence of the growth rate of the national income. In the market-oriented developing countries of today, however, income tends to be concentrated in the hands of a small number of rich people who either do not save significantly on a current basis, or use

4. See footnote 2 above.
their savings for purposes other than domestic investment. In addition, the disposable income of the rich is spent mostly on imports and on goods produced domestically by means of relatively capital-intensive techniques. Thus it is argued that a redistribution of income from the rich to the poor would cause a substantial increase in the demand for those goods and services which have a relatively high labour content, and low capital and import contents. Such a change in the composition of demand would not only help to solve the problems of unemployment and balance of payments disequilibrium, but it would do so without necessarily reducing the growth rate of output in the economy.

The basic idea behind "redistribution with growth" can be analyzed in terms of the two-sector approach adopted in the previous chapters. For this purpose it would seem appropriate either to invoke Eckaus's theory of the "factor-proportions problem", or to turn to the "market imperfections hypothesis" discussed in Chapter 3: in each case we were able to derive a transformation curve partly consisting of a labour-unemployment segment, along which each output combination corresponded to a different rate of labour unemployment. Although Eckaus's theory will be used below for expositional purposes only, it is perhaps worth noting that the conclusions of our analysis are generally applicable also to the market imperfections hypothesis.

---

5. See, for example, ILO (1970).

6. This enables us to avoid the assumption of differential wage rates embodied in the market imperfections hypothesis.
Consider, for example, Figure 5.1 which reproduces part of the information contained in Figure 2.5. The initial equilibrium is given by the output combination, E. Since point E lies on segment TF of the transformation curve, it represents a situation of labour unemployment in the economy as a whole. Similarly, the fact that the commodity price line has the same slope as TF means that the optimal distribution point occurs at E' on the contract curve, O_aE.² The community indifference curve labelled Sₗ represents a so-called Scitovsky curve (rather than a Bergson frontier) insofar as it depends in part on the distribution of income between two individuals, A and B:³ specifically, Sₗ corresponds to the particular distribution given by point E' on the contract curve, O_aE.

Let us now introduce a change in the personal distribution (or ownership) of the production factors, and hence in that of income.⁹ Suppose income is redistributed from individual B to individual A, so that B now demands less of both X and Y while A in turn demands more of both. But since A has a stronger relative preference for good X than does B, the increase in A's demand for X will be greater than the decrease in B's demand for the same good; and similarly, the decrease in B's demand for good Y will exceed the increase in A's corresponding

---

7. That is, price line ee" is parallel to TF.
8. See Graaff (1957, ch. 3), Mishan (1960) and the references listed in Appendix 1.
demand. The net effect is thus an increase in the aggregate demand for good X, and a decrease in demand for Y, *ceteris paribus*, raising the demand price of X relative to that of Y; that is, the redistribution of income gives rise to a rightward shift in the market demand function for good X, and a leftward shift in the demand function for Y. This may be illustrated diagrammatically with the aid of Figure 5.1. The redistribution of income causes a shift of the distribution point from E' to E", the effect of which is twofold: firstly, the increase in the demand price for X raises the slope of the relative price line from that associated with TF, to that given by the new (demand) price line, mm'; and secondly, it establishes a new Scitovsky curve S_1', which is tangent to mm' at point E. This implies that the new disequilibrium price ratio now exceeds the old equilibrium price ratio (given by the slope of TF); or put differently, the demand price ratio exceeds the (as yet unchanged) supply price ratio. Accordingly, the economy is now experiencing both an excess demand for good X and an excess supply of good Y.

Given the existence of competitive conditions in both markets, these discrepancies will lead to an increase in the quantity supplied of good X, and to a decrease in the quantity supplied of Y, thus reducing the newly established disequilibrium price ratio while at the same time raising the old equilibrium price ratio. These changes will continue until the two price ratios eventually meet at some intermediate point, e.g. F in Figure 5.1, where the new equilibrium price ratio is given by
the slope of price line ff'. At point F the new Scitovsky curve, $S_2$, is tangent both to the transformation curve and to price line ff'; and similarly, the corresponding distribution point occurs at $F'$ on the new contract curve, $O_4F$. Since point F also represents a situation of overall full employment, it follows that the redistribution of income from B to A has effectively removed the labour unemployment.

The latter conclusion derives from the fact that income was redistributed in favour of the individual with the stronger relative preference for the more labour-intensive good. The subsequent increase in demand for the labour-intensive good, coupled with the decrease in demand for the capital-intensive good, accounted for the net increase in labour employment in the economy.

3. Income Redistribution and the Demonstration Effect

The previous example relied on a number of seemingly unrealistic assumptions. Apart from the generally static nature of the analysis, there may be no justification for the assumption that individual tastes would remain constant in the face of a policy of income redistribution. It is quite possible that individual tastes may change over time in accordance with Engel's well-known law, according to which the proportion of income spent on food and other such necessities tends to decline with increases in the level of individual income; in a critique of the notion of declining marginal income utility, for example, Musgrave (1959, p. 103) observed that, "rising
needs develop with rising income". One possible explanation for this tendency is provided by the so-called demonstration effect, a concept used by Pigou (1960, pp. 89-92), Duesenberry (1949) and others in their analyses of the consumption behaviour of poor people. According to these authors, the poor are generally subject to a demonstration effect insofar as they emulate, or try to emulate, the consumption habits of the rich. Should the income of the poor increase as a result of a policy of redistribution, for example, their spending pattern would be likely to change in such a way as to reflect more closely that of the rich.

The demonstration effect has gained in importance recently due to the growth of international trade and the accompanying improvements in communication and transportation networks within the developing countries and between the developed and developing countries. These improvements have facilitated both the flow of information and the transportation of goods and services between the rich and poor regions of the world. Likewise, the expanding activities of multinational companies have brought about a "diffusion of the 'ideology' of the consuming society - that is, the creation of international consumption habits and standards, through the product differentiation and advertising characteristic of oligopolistic industries, throughout the developed and underdeveloped world".

10. See, for example, Nurske (1953, ch. 3) and Meier and Baldwin (1963, pp. 308-310).
11. Their main proposition was that the magnitude of household saving depends not on the absolute level of household income, but rather on the household income relative to that of other households.
The emergence of a demonstration effect may limit the employment potential of a policy of income redistribution. Consider, for example, Figure 5.2 which reproduces the situation shown in Figure 5.1: points F and F' represent the equilibrium obtained after income has been redistributed from B to A, on the assumption that individual tastes have remained unchanged. Suppose, however, that A is subject to a demonstration effect in the sense that his taste depends in part on B's consumption pattern; specifically, let the increase in A's income *ceteris paribus* raise his relative preference for good Y. Such a situation is shown in Figure 5.2 by the anti-clockwise rotation of A's indifference curves, for example from A₂ to A₂' along price line ff" passing through point F'. This implies, in turn, an increase in the aggregate demand for Y which, *ceteris paribus*, raises its demand price relative to that of good X - indicated, for example, by the difference between the slope of the new price line, nn', and that of the previous price line ff' at point F. Given overall competitive conditions, however, the resultant increase in the supply of good Y will reduce its temporarily prevailing (disequilibrium) price, until equilibrium is re-established at a point such as G. At G the indifference curve labelled S₃ is tangent to the new (equilibrium) price line, TF, which coincides with the labour-unemployment segment of the transformation curve; similarly, the corresponding (optimal) distribution point is given by G' on the new contract curve OₐG. Since point G lies on the labour-unemployed segment of the transformation
curve, however, it follows that the demonstration effect generally, and the change in A's taste in particular, have effectively given rise to the emergence of labour unemployment in the economy.

The above example illustrates that the existence of a demonstration effect may limit the employment potential of a policy of income redistribution. This would occur if the policy were accompanied by an increase in the income recipient's relative preference for the capital intensive good. Moreover, it is at least possible that the latter change in tastes could cause a net increase in the relative supply of the capital-intensive good; and if initially the economy experienced labour unemployment, for example due to the existence of fixed factor proportions, the net effect of the policy would then be to increase the level of labour unemployment. Generally, it is clear that the policy-induced increase in demand for the capital-intensive good would, ceteris paribus, raise the quantity of capital needed to attain overall full employment in the economy. Such a conclusion would seem to contradict Todaro's (1971, p. 396) (and others') recent rejection of the ".... conventional wisdom of economic theory, which placed top priority on the rapid accumulation of capital as the key to successful economic progress".

4. Income Redistribution and Economic Growth

Turning now to the implications of income redistribution for economic growth, it would seem appropriate to relate Kaldor's (1955-6; 1960) theory of distribution to the fixed-proportions
model outlined in Chapter 2. In contrast to the neoclassical theory of economic growth for, example, Kaldor assumes that the aggregate savings ratio in the economy depends on the distribution of income between capitalists and workers. He divides the national income \( (Y) \) into two components, namely, wages \( (W) \) and profits \( (P) \) which accrue to workers and capitalists respectively. While capitalists' savings are proportionately larger than those of workers, both savings propensities are assumed to be constant. We thus have the following set of equations:

\[
Y = W + P \quad (5.1)
\]

\[
S_w = s_w W \quad (5.2)
\]

and

\[
S_p = s_p P \quad (5.3)
\]

where \( S_w \) and \( S_p \) represent aggregate savings out of wages and profits, respectively, \( s_w \) and \( s_p \) are the corresponding marginal (and average) propensities to save, and \( s_w < s_p \). Total savings are given by

\[
S = S_w + S_p = s_w Y + s_p P \quad (5.4)
\]

or substituting for \( W \) from (5.1),

\[
S = s_w (Y - P) + s_p P = (s_p - s_w)P + s_w Y \quad (5.4')
\]

Dividing (5.4') by \( Y \) gives

\[
S/Y = s = (s_p - s_w)(P/Y) + s_w \quad (5.5)
\]
Equation (5.5) states that the savings propensity for the economy as a whole, $s$, is a positive function of the share of profit in the national income, $P/Y$.

The latter relationship is particularly relevant to the factor-proportions problem discussed in Chapter 2. Recall that the growth rate of output for the economy characterized by fixed factor proportions, is given by

$$y = k = \ell$$

(5.6)

where $k$ and $\ell$ are the growth rates of capital and labour employment, respectively. Similarly,

$$k = s\sigma$$

(5.7)

where $\sigma$ is the constant output/capital ratio. Substituting (5.5) into (5.7) gives

$$k = \sigma (s_p - s_w)(P/Y) + s_w$$

(5.7')

which states that the growth rate of capital, and hence the growth rates of output and labour employment, are directly related to the share of profit in the national income.

Suppose now that the growth rate of labour unemployment is given by $u = l_a - \ell$, where $l_a$ is the (constant) growth rate of the labour supply, and $l_a > \ell$. According to equation (5.7'), it would be possible to eliminate such unemployment by raising the level of aggregate savings via a redistribution of income in favour of the capitalists; specifically, the increase in $s$ should be such as to raise both $k$ and $\ell$ by the proportion $u$. 
so that $k' = l' = l_a$ where $k' = k + u$ and $l' = l + u$.

Conversely, a redistribution of income in favour of workers would reduce the growth rates of capital and labour employment, and if the labour supply grew at a constant rate, the net effect would then be an increase in the level of labour unemployment.

The above analysis suggests that a conflict may exist between the objective of a high growth rate of output on the one hand, and that of an equitable distribution of income on the other. The extent of this conflict will, of course, depend on whether the propensity to save varies significantly between different income groups within the economy, and/or whether aggregate savings do indeed contribute markedly to the growth of output. Generally, if the savings propensity of the rich exceeded that of the poor, a policy of income redistribution in favour of the poor would *ceteris paribus* cause a decrease in the aggregate savings level, and hence also in the growth rates of capital, output and employment.

There are good reasons why the latter conclusion should be treated with caution. It fails to recognize that people earning high incomes may use their savings for investment purposes in other countries;¹⁴ or alternatively, that private domestic investment may not account for all or even most of the secular growth of income.¹⁵ Moreover, the analysis is based on

---

¹⁴. See ILO (1970).

¹⁵. See Meier (1976, pp. 258-261).
a given state of technology which may well be subject to change over time: there is enough evidence indicating that the labour content of production can be increased substantially by the use of more appropriate production techniques. \(^{16}\) Nevertheless, even if a redistribution of income in favour of the poor did cause a decrease in the growth of output, the policy might still be deemed desirable from a social welfare point of view: "... it is worth sacrificing production to reduce this evil (of unemployment)". \(^{17}\)

\(^{16}\) See, for example, Pack (1976).

\(^{17}\) Stewart and Streeten (1973, p. 372).
CONCLUSION

One of the chief objectives of this thesis was to investigate whether 'conventional' economic theory - specifically the neoclassical theory of general equilibrium - is sufficiently flexible to accommodate the particular conditions prevailing in the developing countries. Most existing theories of economic underdevelopment adopt an analytical approach which in effect amounts to relaxing some of the basic assumptions of the neoclassical theory. When applied to the two-sector model of general equilibrium, these theories generally yield predictions which are vastly different from those associated with the standard neoclassical assumptions of perfect competition, unlimited factor substitutability and unrestricted resource mobility.

Several theories attempt to explain the development problem in terms of the specific production processes used in developing countries. Myrdal's (1957) theory of cumulative causation, for example, effectively introduces increasing returns to scale in at least one sector or region of the economy; in contrast to the neoclassical theory, he thus envisages a cumulative process of regional divergence in the income level per worker. Similarly, Richard Eckaus's (1955) explanation of the "factor-proportions problem" is based on the assumption of limited factor substitutability. This enables him to establish the existence of a so-called "unemployment equilibrium", thus implying that developing countries may be faced with a conflict between the objective of
maximizing social welfare on the one hand, and that of full employment on the other. More recently, Leibenstein (1960) has shown how this trade-off could be complicated by the introduction of labour-saving technological inventions and innovations. The solution to the factor-proportions problem generally consists in the adoption of more appropriate, usually labour-biased technologies, increased capital formation and a reduction in the growth rate of the labour supply.

Much of the postwar literature on economic development centres on the imperfectly competitive structure of the product and the factor markets in developing countries. Myint (1954) has highlighted the role played by monopolies and oligopolies during the so-called "opening-up" process of economic development. Likewise, both Lewis's (1954) dualist theory and Todaro's (1969; 1971) model of rural-urban migration attempt to explain the unemployment problem in terms of various factor price distortions. Prebisch (1959) and Singer (1950) have again shown how prevailing differences in the structure of markets between the developed and developing countries might turn the international terms of trade against the latter; using a two-sector model, Bhagwati (1958) has demonstrated that such a deterioration in the terms of trade could bring about a net decrease in the welfare level of the countries concerned. Generally, the policy measures relevant to the "market imperfections" problem include the creation of job opportunities in the rural (rather than urban) sector, the encouragement of informal-sector enterprises, and the imposition of appropriate
factor taxes and subsidies as a means of counteracting the adverse effect on employment of factor price distortions.

A more recent approach to the unemployment problem is the plea by the International Labor Office (1972) for a redistribution of income within the developing countries. In terms of the two-sector model, such a policy may well succeed in eliminating labour unemployment caused by fixed factor proportions and/or factor price distortions. It should be realized, though, that a redistribution of income might also reduce the aggregate savings level, and hence the growth rates of capital and labour employment in the economy.

On the whole, it would seem that these theories do indeed adopt a modified version of the neoclassical theory in providing a fairly comprehensive explanation of the economic problems of labour unemployment, low incomes and inequality; that is, the theories generally seek to explain the development problem in terms of the operation of the input and output markets, as well as in relation to the production process itself. However, a potentially damaging criticism would be that most of these theories are too narrowly based on the so-called "economic" determinants of the problem.¹ This raises the question whether particular environmental influences, social institutions and individual attitudes are sufficient to render the assumption of "optimizing behaviour" largely inappropriate within the context of the developing nations.² While the evidence on this is far from conclusive, there does seem to be a growing

¹. See, for example, Myrdal (1968, 1973) and Seers (1963).
². See Boeke (1953).
consensus among development economists at least, that the majority of people in developing countries do respond positively to a wide range of market incentives.\(^3\) It is indeed worth quoting Baldwin's (1972) optimistic view on this matter: "Since economic development via reliance upon the price system is the most inexpensive way for governments to carry out their development commitments, the mass of evidence indicating that most people respond favourably to economic incentives is a very hopeful sign".

Mainstream development economists do not, of course, deny that some "non-economic" variables may play an important role in restraining individual responses to market signals.\(^4\) What is being questioned increasingly, however, is the feasibility of systematically treating these "non-economic" variables in such a way that they may be incorporated into an explanatory model of the development problem. Indeed, the literature seems to suggest that such a procedure is either not possible at the present time, or that it could only be adopted at the risk of destroying the operational value of the model(s) concerned.

Of course, the meaningful coexistence of 'theory' and 'practice' should not require a denial of their respective essential and distinctive features. A tolerant attitude towards such a compromise was adopted by a recent review of

---

3. See, for example, Neumark (1958), Elkan (1960), Lipton (1968) and Baldwin (1972).
Sir John Hicks' 'Collected Essays in Economic Theory':

"..... Sir John shows that it is pretty well impossible for practice ever to be more than a kind of make-do and mend. However, there is little point in rejecting the real world because it is not the ideal world of an abstract theoretical system (although that is precisely what some economists have done). Instead, it is necessary to make sense of the real world, but to retain the intellectual standards of the theoretical one - as Sir John always does".5


The presumed existence of a social welfare function implies the following important suppositions:

(i) The welfare of society depends only on the welfare of the individuals comprising that society; that is,

\[ W = W(U_a, U_b) \]  \hspace{2cm} (A1.1)

where \( W \) represents the level of social welfare, and \( U_a \) and \( U_b \) are the utility levels of individuals A and B respectively.

(ii)(a) The social welfare function presupposes an ethical valuation of the relative deservingness or worthiness of the individuals concerned.\(^1\) Consider, for example, the following social or community indifference curve,

\[ W_0 = W(U_a, U_b) \]  \hspace{2cm} (A1.2)

Its rate of change,

\[ (\frac{aW}{aU_a}) \frac{dU_a}{aW} + (\frac{aW}{aU_b}) \frac{dU_b}{aW} = 0 \]  \hspace{2cm} (A1.3)

or

\[ \frac{dU_a}{dU_b} = -\frac{\frac{aW}{aU_b}}{\frac{aW}{aU_a}} \]  \hspace{2cm} (A1.3')

reflects society's relative preference for the welfare of the two individuals: that is, \( (dU_a/dU_b) \) depends on the relative contributions assumed to be made by the respective individual utilities to the level of social welfare.

---

1. Irrespective, that is, of whether such a valuation is made democratically or dictatorially, and whether or not it is "non-controversial". (See Nath (1973, pp. 20-21) and Arrow (1951)).
(ii)(b) The above value judgement is embodied in what is sometimes referred to as the Bergson welfare criterion.² This criterion encompasses, or "is more complete than",³ the so-called Pareto, Kaldor-Hicks, Scitovsky and other welfare criteria.⁴ Specifically, although the Bergson welfare criterion is consistent with the Pareto criterion — for example, social welfare will improve *aeteris paribus* if the utility level of at least one individual increases; or $(\delta W/\delta U_a) > 0$ — it is nevertheless compatible also with the possibility of a welfare improvement arising from an increase in the utility level of one individual at the expense of that of the other.

(iii) The individual (or head of household) is the sole (and best possible) judge of his own welfare; or

$$U_a = U_a (Q^a_x, Q^a_y) \quad (A1.4)$$

and

$$U_b = U_b (Q^b_x, Q^b_y) \quad (A1.5)$$

where $Q^a_x, Q^a_y$ and $Q^b_x, Q^b_y$ are the quantities of $X$ and $Y$ consumed by A and B respectively.

Let us now extend the analysis somewhat, by substituting (A1.4) and (A1.5) into (A1.1):

$$W = W [U_a (Q^a_x, Q^a_y), U_b (Q^b_x, Q^b_y)] \quad (A1.1')$$

2. Bergson (1937-38), Baumol (1965, p. 167) refers to Bergson's criterion as "the social value judgement par excellence".


4. See Kaldor (1939), Hicks (1939), Sitovsky (1941-42) and Little (1957).

5. This implies in effect "complete consumer and producer sovereignty" (Nath, 1969, p.9).
But since,

\[ Q_x = Q_x^a + Q_x^b \]  
\[ Q_y = Q_y^a + Q_y^b \]

(A1.6)

(A1.7)

where \( Q_x \) and \( Q_y \) are the total quantities of \( X \) and \( Y \) respectively; then, using the "function of a function" rule, 6

\[ W = V(Q_x, Q_y) \]
\[ W_0 = V(Q_x, Q_y) \]

(A1.1"

(A1.8)

where (A1.1") represents the social welfare function in commodity space, and (A1.8) is one of the (infinite number of) corresponding community indifference curves. Two such community indifference curves are shown in Figure A1.1, where it is assumed that output ratios along the curve labelled \( W_2 \) all yield a higher level of social welfare than those given by \( W_1 \).

Each of the curves in Figure A1.1 may be viewed as one of a set of non-intersecting Scitovsky community indifference curves, 7 provided it is assumed that individual tastes and/or marginal propensities to consume are identical. 8 Should such an assumption seem unnecessarily far-fetched, then an alternative interpretation is provided in the form of the so-called Bergsen welfare frontier. 9

6. A similar procedure is followed by Winch (1971, p.50).
7. Their derivation is illustrated in Mishan (1960) and Bilas (1972, ch.12).
8. See Gorman (1953) and Nath (1969, p.27).
9. This was introduced in the literature by Samuelson (1956) and Graaff (1957, ch.3).
FIGURE A1.1
A Bergson frontier represents the inner limit of a set of intersecting Scitovsky indifference curves, each of which is associated with given levels of individual utility, and all of which correspond to a particular level of social welfare: in short, each point along the Bergson frontier represents the maximum attainable welfare limit, on the assumption that the goods are optimally produced and distributed among the members of society. Since Bergson frontiers cannot intersect in the absence of externalities, they represent in effect a set of Pareto-comparable community indifference curves in output space.

10. See Graaff (1957, ch.3)
APPENDIX 2:  Welfare Maximization and Pareto Optimality.

Consider the following Lagrangian expression,

\[ G = W \left[ U_a(Q_x^a, Q_y) + U_b(Q_x^b, Q_y^b) \right] \\
+ P_x \left[ Q_x(K_x, L_x) - Q_x^a - Q_x^b \right] + P_y \left[ Q_y(K_y, L_y) - Q_y^a - Q_y^b \right] \\
+ r (K_x + K_y - \bar{K}) + w (L_x + L_y - \bar{L}) \]  \hspace{1cm} (A2.1)

where the variables are defined as in the text, and \( P_x, P_y, r \) and \( w \) are Lagrange multipliers representing the prices of \( X, Y, K \) and \( L \) respectively. Maximization of (A2.1) implies that:

\[ \frac{\partial G}{\partial Q_x^a} = \left( \frac{\partial W}{\partial U_a} \right) \left( \frac{\partial U_a}{\partial Q_x^a} \right) - P_x = 0 \]  \hspace{1cm} (A2.2)

\[ \frac{\partial G}{\partial Q_y^a} = \left( \frac{\partial W}{\partial U_a} \right) \left( \frac{\partial U_a}{\partial Q_y^a} \right) - P_y = 0 \]  \hspace{1cm} (A2.3)

\[ \frac{\partial G}{\partial Q_x^b} = \left( \frac{\partial W}{\partial U_b} \right) \left( \frac{\partial U_b}{\partial Q_x^b} \right) - P_x = 0 \]  \hspace{1cm} (A2.4)

\[ \frac{\partial G}{\partial Q_y^b} = \left( \frac{\partial W}{\partial U_b} \right) \left( \frac{\partial U_b}{\partial Q_y^b} \right) - P_y = 0 \]  \hspace{1cm} (A2.5)

\[ \frac{\partial G}{\partial K_x} = P_x \left( \frac{\partial Q_x}{\partial K_x} \right) + r = 0 \]  \hspace{1cm} (A2.6)

\[ \frac{\partial G}{\partial L_x} = P_x \left( \frac{\partial Q_x}{\partial L_x} \right) + w = 0 \]  \hspace{1cm} (A2.7)

\[ \frac{\partial G}{\partial K_y} = P_y \left( \frac{\partial Q_y}{\partial K_y} \right) + r = 0 \]  \hspace{1cm} (A2.8)

\[ \frac{\partial G}{\partial L_y} = P_y \left( \frac{\partial Q_y}{\partial L_y} \right) + w = 0 \]  \hspace{1cm} (A2.9)

The results in the text are obtained as follows:

(i) Rearranging equations (A2.6) through (A2.9) and dividing (A2.7) by (A2.6) and (A2.9) by (A2.8), we get:

\[ \frac{\partial Q_x}{\partial K_x} = \frac{w}{r} \frac{\partial Q_y}{\partial K_y} \]  \hspace{1cm} (A2.10)
(ii) Rearranging equations (A2.2) through (A2.5) and dividing (A2.2) by (A2.3) and (A2.4) by (A2.5), give:

\[ \frac{aU/aQ_x}{aU/aQ_y} = \frac{P_x}{P_y} \ \
\frac{aU/aQ_x}{aU/aQ_y} = \frac{aU/aQ_y}{aU/aQ_x} \]  

(A2.11)

(iii) Next, rearranging (A2.6) and (A2.8) gives:

\[ P_x \left( \frac{aQ_x/aK_x}{aQ_x/aK_y} \right) = P_y \left( \frac{aQ_y/aK_y}{aQ_y/aK_y} \right) \]  

(A2.12)

or

\[ \frac{aQ_y/aK_y}{aQ_x/aK_x} = \frac{P_x}{P_y} \]  

(A2.12')

The total differentials of the two production functions are:

\[ dQ_y = \left( \frac{aQ_y/aK_y}{aQ_x/aK_x} \right) dK_y + \left( \frac{aQ_y/aL_y}{aQ_x/aL_y} \right) dL_y \]  

(A2.13)

\[ = \left( \frac{aQ_y/aK_y}{aQ_x/aK_x} \right) dK_y \left[ 1 + \frac{\left( \frac{aQ_y/aL_y}{aQ_x/aK_y} \right) dL_y}{\left( \frac{aQ_y/aK_y}{aQ_x/aK_x} \right) dK_y} \right] \]  

(A2.13')

and similarly,

\[ dQ_x = \left( \frac{aQ_x/aK_x}{aQ_x/aK_x} \right) dK_x \left[ 1 + \frac{\left( \frac{aQ_x/aL_x}{aQ_x/aK_x} \right) dL_x}{\left( \frac{aQ_x/aK_x}{aQ_x/aK_x} \right) dK_x} \right] \]  

(A2.14)

But since \( dK_x = -dK_y \) and \( dL_x = -dL_y \) under perfect competition, it follows from (A2.10) that the second bracketed term of (A2.13') is equal to that of (A2.14); accordingly, dividing (A2.13') by (A2.14) we get:

\[ \frac{dQ_y}{dQ_x} = \frac{aQ_y/aK_y}{aQ_x/aK_x} \]  

(A2.15)
Combining (A2.12') and (A2.15) gives:

\[ \frac{dQ_y}{dQ_x} = \frac{P_x}{P_y} \quad \text{(A2.16)} \]

and similarly, combining (A2.11) and (A2.16):

\[ \frac{aU_a/aQ_a^x}{aU_a/aQ_a^y} = \frac{aU_b/aQ_b^x}{aU_b/aQ_b^y} = \frac{dQ_y}{dQ_x} \quad \text{(A2.17)} \]

Finally, from (A2.2) through (A2.5) we have:

\[ \frac{aW}{aQ_a^x} = P_x = \frac{aW}{aQ_a^b} \quad \text{(A2.18)} \]

and

\[ \frac{aW}{aQ_a^y} = P_y = \frac{aW}{aQ_a^b} \quad \text{(A2.19)} \]

from which it follows that

\[ \frac{aW}{aQ_x} = \frac{P_x}{P_y} \quad \text{(A2.20)} \]

Combining (A2.16), (A2.17) and (A2.20) gives:

\[ \frac{aW/aQ_x}{aW/aQ_y} = \frac{aU_a/aQ_a^x}{aU_a/aQ_a^y} = \frac{aU_b/aQ_b^x}{aU_b/aQ_b^y} = \frac{dQ_y}{dQ_x} \quad \text{(A2.21)} \]

representing all the necessary and sufficient conditions for the maximization of social welfare.

See footnote 6 in Appendix 1 above.
APPENDIX 3: Factor Mobility and the Neoclassical Steady-State.

The growth rate of the total or combined regional output is given by:

\[ q'_t = \left( \frac{q'_x Q_x + q'_y Q_y}{Q_x + Q_y} \right) \]  \hspace{1cm} (A3.1)

where \( q'_x \) and \( q'_y \) represent the regional growth rates of output under steady-state conditions; that is,

\[ q'_x = n + l_{yx} \]  \hspace{1cm} (A3.2)

\[ q'_y = n + l_{xy} \]  \hspace{1cm} (A3.3)

Substituting (A3.3) into (A3.2) into (A3.1), and rearranging terms we have,

\[ q'_t = n + \left[ h \left( bK_x^* - fK_y^* \right) \left( \frac{Q_x}{L_x} - \frac{Q_y}{L_y} \right) \right] / (Q_x + Q_y) \]  \hspace{1cm} (A3.1')

But since output per worker is assumed to be a constant proportionate function of the capital: labour ratio, or

\[ \frac{Q_x}{L_x} = bK_x^* / (1 - \alpha) \]  \hspace{1cm} (A3.4)

and \( \frac{Q_y}{L_y} = fK_y^* / (1 - \beta) \)

substitution of (A3.4) into (A3.1') gives:

\[ q'_t = n + \left[ h(bK_x^* - fK_y^*) \left( \frac{bK_x^*/(1-\alpha) - fK_y^*/(1-\beta)}{Q_x + Q_y} \right) \right] \]  \hspace{1cm} (A3.1'')

Accordingly, since \( bK_x^* < fK_y^* \) and \( (1 - \alpha) > (1 - \beta) \) by assumption, it follows that \( q'_t > n \) and \( (q'_t - n) > 0 \).


Bairoch, P. (1975), The Economic Development of the Third World since 1900, Methuen.


______ (1966), Causes of the Slow Rate of Growth of the United Kingdom, Inaugural Lecture, Cambridge Univ. Press.


Pareto, V. (1897), Cours d' économie politique II, Rome.


Verdoorn, P.J. (1947), "Fattori Che Regolano Lo Svillupo Della Produttività Del Lavoro", *L' Industria*.


